The Hard Problem of Intertheoretic Comparisons

According to *metanormativists*,[1] moral uncertainty can affect how we *ought*, in some morally significant sense, to act. The "ought" associated with moral uncertainty is a "subjective 'ought' ": it is sensitive to the agent's limited information, and therefore to her (rational) uncertainty. The traditional subjective "ought" is sensitive only to the agent's *descriptive* uncertainty. The metanormativist holds that *normative* uncertainty also affects how an agent ought to act: what choices are (ir)rational, (un)justified, morally praise(blame)worthy, or (in)appropriate.[2]

Metanormativism is a controversial view. The aim of this paper is not to provide a positive argument for metanormativism, but to show how it can avoid a problem that is sometimes taken to be insuperable.[3] The traditional model for rational decision-making under uncertainty, expected choiceworthiness[4] maximization, makes use of quantitative choiceworthiness assignments, the expectations of which determine which options are permissible. In generalizing to moral uncertainty, metanormativists represent distinct moral theories with choiceworthiness functions. This leads to what I'll call the **easy problem of intertheoretic comparisons**: the worry that there's no way to compare the choiceworthiness assignments of one theory to those of another. Metanormativists have propounded various solutions to the easy problem, surveyed in §1; while none has become orthodox, there is reason for optimism.

But another worry looms. While we might be able to represent some moral theories with cardinal choiceworthiness functions, others are thought to be *merely ordinal*. Such theories qualitatively rank outcomes, but provide no quantitative information about the ranking. Any cardinal choiceworthiness function would misrepresent such theories. It would present an ordinal theory as opinionated in places where the theory genuinely was not. Whenever an agent assigns positive credence to such a theory, expected choiceworthiness maximization becomes unusable. Merely ordinal choiceworthiness can't figure into an *average*. I call this the *hard problem of intertheoretic comparisons*.[5]

[1] I borrow this term from MacAskill (2014); Harman (2011) calls this position "uncertaintism."

[2] Metanormativists differ on these finer points.

[3] For discussion, see e.g. Hedden (2016); MacAskill (2016).

[4] I use "choiceworthiness" as a generalization of "utility" or "value", both to avoid the presupposition of utilitarianism or consequentialism and to accommodate the possibility that the relevant quantity may be "subjective", in the sense of the subjective "ought".

[5] Metanormative decision theories face other problems that this paper doesn't address, e.g., problems associated with absolutist theories that apparently assign infinite degrees of choiceworthiness, and with normative uncertainty about metanormative decision theory itself.

In §2, I argue that to solve the hard problem, we should model moral theories, and moral hypotheses in general, with imprecise choiceworthiness. But generalizing familiar decision theories for imprecise choiceworthiness to the case of uncertainty between moral theories generates puzzles: it seems to require reifying parts of the model that don't correspond to anything in moral reality. §3 examines three ways of addressing this problem: by demystifying the reified elements by using them as promiscuously as possible; by constructing alternative decision theories that don't require the troublesome elements; and by employing an alternative model of metanormative decision problems, and of moral uncertainty more generally.

1 Moral uncertainty and the easy problem

1.1 Moral uncertainty

How should we make decisions when we're uncertain about what will result from our actions? The traditional answer in decision theory is that rational agents *maximize expected choiceworthiness.* A natural hypothesis is that this proposal generalizes to cases where agents are uncertain not merely about what outcomes will result from their acts, but also about how morally good or bad specific outcomes are. Versions of this have been argued by Oddie (1994); Lockhart (2000); Ross (2006); MacAskill (2014); Tarsney (2018); Carr (2020).

Expected choiceworthiness is determined by a *probability function* and a *choiceworthiness function* (CF). In traditional decision theories, the probability function represents a rational agent's degrees of belief, while the CF represents the same agent's degrees of desire.[6] Each option available to an agent is assigned an *expected choiceworthiness*: a probability-weighted average of the choiceworthiness of its possible outcomes. Let u be the relevant CF, cr the relevant subjective probability function, $\mathcal{S}$ a partition of possible states of the world, and "$cr(s \mid\mid a)$" the subjective probability of s on the supposition that the agent performs a.[7] Then the expected utility of an act a is:

$$Eu(a) = \sum_{s \in \mathcal{S}} cr(s \mid\mid a)u(s \wedge a)$$

Expected choiceworthiness maximization requires agents to choose options that maximize this quantity.

How can we tweak expected choiceworthiness maximization to accommodate moral uncertainty? Uncertainty about the moral is standardly represented as uncer-

[6] More accurately, non-instrumental, comparative desire; see Phillips-Brown (forthcoming).

[7] This representation is meant to be neutral between causal and evidential decision theory.

tainty over competing moral theories.[8] We'll assume each theory can be represented with a CF. For a moral theory t in the set of moral theories $\mathcal{T}$, let u_t be t's CF. For simplicity, throughout, we'll focus on cases where agents are certain of all relevant descriptive facts. Then we can generalize expected choiceworthiness to expected *intertheoretic* choiceworthiness:[9]

$$Eiu(a) = \sum_{s \in \mathcal{S}, t \in \mathcal{T}} cr(s \wedge t \mid\mid a) \cdot u_t(s \wedge a)$$

If expected choiceworthiness maximization is the correct metanormative decision theory, then an agent is blameworthy, or acts inappropriately, if she fails to maximize this quantity. Other decision rules for descriptive uncertainty can be generalized for moral uncertainty along similar lines.

1.2 The easy problem of intertheoretic comparisons

1.2.1 What CFs represent

The most widely-discussed challenge to the metanormative generalization of expected choiceworthiness maximization is what's often called "the problem of intertheoretic comparisons." The problem arises from how we quantitatively represent choiceworthiness. Choiceworthiness is generally understood to have no upper or lower bounds. Things can always get better or worse; indeed, they can always get *a lot* better or *a lot* worse. It's also traditionally assumed that there's no privileged zero-point for choiceworthiness, where positive choiceworthiness represents goodness or value, and negative choiceworthiness represents badness or disvalue.

These assumptions provide a lot of freedom in assigning a *scale* to choiceworthiness. Any CF is *informationally equivalent*—equivalent in the information it accurately represents about a person or theory's values—to infinitely many other CFs.

Some properties of a CF are nonarbitrary. For example, *order*: once we stipulate that higher quantities are better, the choiceworthiness assigned to x must be greater than the choiceworthiness assigned to y only if x is better than y. Another example: it's traditionally assumed that *ratios of differences* are nonarbitrary: if $\frac{u(x_1)-u(x_2)}{u(x_3)-u(x_4)} = r$, then any CF v that's informationally equivalent to u must preserve that same ratio. In other words: one CF u is informationally equivalent to another v iff there's a positive affine transformation θ such that $\theta(u(x)) = v(x)$ for any possibility x.

[8] Lockhart (2000); Ross (2006); Sepielli (2009), etc. See Carr (2020) for discussion and alternatives.

[9] Alternatively, we may assume that the CFs in the metanormative decision theories throughout assign *subjective* choiceworthiness, already accounting for the relevant state of descriptive uncertainty; see Carr (2020) for discussion. The benefit of this maneuver is that it can incorporate normative uncertainty about (descriptive) decision theory.

The arbitrariness of particular CFs generates no problems for representing an individual person or theory. But it does cause problems for *comparing* choiceworthiness assignments across people or theories. It's generally accepted within decision theory that some kinds of interpersonal choiceworthiness comparisons are meaningless:

- *comparison of magnitudes:* $u(x) \geq v(y)$
- *comparison of units:* $u(x_1) - u(x_0) \geq v(y_1) - v(y_0)$
- *comparison of ratios:* $\dfrac{u(x_1)}{u(x_0)} \geq \dfrac{v(x_1)}{v(x_0)}$

Many have held that the corresponding *intertheoretic* choiceworthiness comparisons are equally meaningless.

More generally, different theories represent different ways of seeing value or rightness in the world. Suppose t_1 holds that the only genuine source of value is happiness, while t_2 holds that the only genuine source of value is cultivating virtues. Neither of these theories weighs the two against each other: on t_1, cultivating virtues has no intrinsic value, and on t_2, happiness has no intrinsic value. So on each theory, there is no exchange rate in the choiceworthiness of happiness and virtue; only one has any value, or any effect on choiceworthiness, whatsoever.

If degrees of choiceworthiness are not comparable across theories, however, then there can be no meaningful assignment of expected intertheoretic choiceworthiness. We can't average incomparable quantities, any more than we can average 9 inches with 14 degrees Celsius. This is what I call **the easy problem of intertheoretic comparisons**.

1.2.2 Solving the easy problem

Many philosophers are optimistic about solving the easy problem of intertheoretic comparisons.

On some accounts, intertheoretic comparisons are grounded in *intuitive agreements* between theories. Ross (2006), for example, resolves the problem by finding some possibilities i and j that aren't points of contention between the two theories, and defining the difference in choiceworthiness between i and j according to t_1 as a unit of choiceworthiness. Then t_2's CF is scaled up or down so that $u_{t_1}(i) - u_{t_1}(j) \geq u_{t_2}(i) - u_{t_2}(j)$, allowing intertheoretic comparisons of units. Tarsney (2018) defends a related account: competing theories sometimes agree about some sources of contributory value. (For example, two theories might agree in how the evaluate pleasure as a source of noninstrumental value even if one and not the other treats beauty as a second source of noninstrumental value.) We then scale the theories' CFs to match with respect to these contributory values, which can provide a basis for comparing their evaluations of possibilities where the theories disagree.

Other accounts ground intertheoretic comparisons in *structural agreements* between theories. Lockhart (2000) offers such a proposal: in each decision problem, each theory determines a maximum and minimum attainable degree of choiceworthiness for the agent. We can then normalize these to generate local intertheoretic comparisons.[10] Sepielli (2009) offers another: if t_1 and t_2 agree about a particular ratio of choiceworthiness differences, this can be used to generate comparisons of units. More precisely: if $\frac{u_{t_1}(x)-u_{t_1}(y)}{u_{t_1}(y)-u_{t_1}(z)} = \frac{u_{t_2}(x)-u_{t_2}(y)}{u_{t_2}(y)-u_{t_2}(z)}$, then $u_{t_1}(x)-u_{t_1}(y) = u_{t_2}(x)-u_{t_2}(y)$.[11]

MacAskill (2014) argues that we should understand intertheoretic choiceworthiness comparisons on the model of absolutism about physical quantities like electric charge. While the choice of scale for measurements of electric charge (coulomb, faraday…) is arbitrary up to positive linear transformation, there are objective, intrinsic charge properties of objects that ground comparisons like 'x has greater charge than y'. So similarly, MacAskill argues, there are objective, intrinsic choiceworthiness properties that are represented by choiceworthiness assignments and that ground intertheoretic comparisons.

Finally, Carr (2020) generalizes expected choiceworthiness maximization for normative uncertainty without allusion to theories, and therefore without the need for intertheoretic choiceworthiness comparisons. On this view, we ought to maximize expected *de dicto* choiceworthiness, under the description "the genuine degree of moral choiceworthiness" of options (where the possible choiceworthinesses of options are locally defined relative to each other). Where Λ is a finite set containing all possible choiceworthiness assignments for a, and "u" is a definite description ("the genuine moral CF, whatever it is"), we can define the *de dicto* expected choiceworthiness of a as: $E\mathrm{u}(a) = \sum_{\lambda \in \Lambda} cr(\mathrm{u}(a) = \lambda) \cdot \lambda$.

Each of these approaches is subject to objections, even significant ones. But they give some reason for optimism about the easy problem. I'm partial to the Carr (2020) approach, and will provide some reasons to favor a generalization of it. But this paper doesn't presuppose it, and will initially focus on approaches to genuinely inter*theoretic* proposals below.

This paper assumes that the easy problem is solvable; we'll focus on a further problem of intertheoretic comparisons. If the easy problem is solvable, I argue, there are good prospects for resolving this harder problem.

[10] But see Ross (2006) and Sepielli (2013) for objections to this proposal.

[11] But see MacAskill (2014) for objections.

2 A harder problem

2.1 Merely ordinal theories

What about agents who give positive credence to merely ordinal theories? How can we formulate decision theories for normative uncertainty between cardinal and ordinal theories?

This is the **hard problem of intertheoretic comparisons.** The challenge is to construct a decision theory that could prescribe rational choices for uncertainty about the measure-theoretic metaphysics of moral choiceworthiness. Such a decision theory would allow for decision problems where some outcomes have cardinal choiceworthiness and some have merely ordinal choiceworthiness:

	t_1	t_2
a_1	$u(a_1) = -83$	a_1
		$\vee$
a_2	$u(a_2) = 23$	a_2

It might be tempting to use decision theories that require only ordinal information: for example, traditional voting procedures.[12] But this amounts to treating all theories as merely ordinal theories. It throws out information that cardinal theories consider important.

It might equally be tempting to assign merely ordinal theories arbitrary cardinal information: for example, if there are n possible totally ordered outcomes, assign the highest ranked outcome choiceworthiness n, the next highest choiceworthiness $n-1$, ...and the lowest ranked outcome choiceworthiness 1.[13] But again, this amounts to treating merely ordinal theories as cardinal theories, with unjustified commitments about their precise cardinal choiceworthiness assignments.

What alternatives are there?

First: we should question the assumption that a rational agent could assign *any* positive credence to the hypothesis that the true moral theory's choiceworthiness assignments have only ordinal structure. Suppose an agent has three options:

1. Donate $5000 to the Against Malaria Foundation

[12] See MacAskill (2016) for discussion of appropriate voting procedures for normative uncertainty across merely ordinal theories.

[13] The strategy ultimately endorsed in (MacAskill, 2016) is arguably an instance of this method. See Martinez (Unpublished) for discussion and objections.

2. Donate \$4999 to the Against Malaria Foundation

3. Bomb a school

A merely ordinal theory could claim that option 2 is worse than option 1, and option 3 is worse than 1 and 2. But it would decline to commit to the claim that while 2 is a little worse than 1, 3 is a lot worse than either. That's absurd. Assigning such a theory positive credence might never be rational.

I won't assume that there are no credible merely ordinal theories. But more plausible, on my view, is the idea that an adequate moral theory might have *less cardinal structure* than our use of CFs traditionally supposes. They might not be fine-grained enough to assign precise numbers to every possibility, or to any. But they are still capable of representing some possibilities not merely as better than others, but as *a lot* better.

On the traditional picture, where CFs are unique up to positive affine transformation, each CF imposes a total order on outcomes: for every pair of outcomes, o_i and o_j, either $u(o_i) \geq u(o_j)$ or $u(o_j) \geq u(o_i)$. But it might be that the correct moral theory doesn't weigh in on the comparative ranking of some outcomes: for some o_i and o_j, neither $u(o_i) \geq u(o_j)$ nor $u(o_j) \geq u(o_i)$ is true. Or for some o_i, o_j, o_k, o_l, it doesn't entail a unique r such that $\frac{u(o_i)-u(o_j)}{u(o_k)-u(o_l)} = r$.

A natural way to represent such hypotheses uses *imprecise choiceworthiness*.

2.2 *Imprecise choiceworthiness*

At first pass, imprecise choiceworthiness assignments represent the degree of choiceworthiness of an outcome not as a single number, but as spread out across multiple numbers. These numbers aren't meant to represent *uncertainty* about choiceworthiness—quantities that might, for all we know, be the degree of choiceworthiness of a possibility. Rather, they represent the range over which a person or theory's assignment of degrees of choiceworthiness are indeterminate. There is no precise degree of choiceworthiness to be uncertain about. An analogy: suppose I asked you for the latitude of Chile. The most accurate response wouldn't be to give a single precise number, but instead a range (17°–56° South).

Imprecise choiceworthiness is analogous to imprecise credence, but characterizing imprecise choiceworthiness requires more care. Precise and imprecise credences are often distinguished as follows: precise credences are represented by a unique credence function and imprecise credences are represented by a set of credence functions (a "credal representor"). One might think: we can represent imprecise choiceworthiness as a set of CFs.

The problem is that even *precise* degrees of choiceworthiness are represented with sets of CFs. This is because the levels, units, and ratios of choiceworthiness assignments are generally regarded as nonsignificant. Let's say that some set of CFs U is *informationally adequate* for an agent or theory x just in case for any proposition p about degrees of choiceworthiness, the elements of U are unanimous about p if and only if p is a genuine commitment of x. Suppose that x's assessments of comparative choiceworthiness satisfy the von Neumann-Morgenstern axioms. Then there's a set of CFs $[u]^{pa}$ that is informationally adequate for x, where all $u' \in [u]^{pa}$ are related by positive affine transformation. Any member of this set is not informationally adequate for x: each carries quantitative information beyond x's body of preferences. We can use precise CFs to model preferences only with the background stipulation that magnitudes, units, and ratios are nonsignificant.

What, then, is the difference between precise and imprecise choiceworthiness? Here are two important differences: precise degrees of choiceworthiness determine a total ordering of outcomes by choiceworthiness. For imprecise choiceworthiness, there are cases where neither of two outcomes is more choiceworthy than the other, nor are they tied. For precise choiceworthiness, there are determinate and precise ratios of choiceworthiness differences; not so for imprecise choiceworthiness.

Suppose a set of choiceworthiness functions U accurately represents x's commitments except that U's information about magnitudes, units, and ratios is nonsignficant. Let Θ be the set of all positive affine transformations. Then the kind of object that's informationally adequate for x will be a "super-representor" $[U]^{pa}$:

$$[U]^{pa} = \{U^* : \exists \theta \in \Theta \, . \, \forall u \in U \, . \, \theta(u) \in U^*\}$$

That is, $[U]^{pa}$ is a set of sets of CFs—a set of representors—each representing imprecise choiceworthiness along different scales (determined by different θs). Call each informationally equivalent $U \in [U]^{pa}$ a "representor" for x. We'll arbitrarily use members of super-representors for the decision theories under discussion, since these theories' verdicts will be invariant across a super-representor's members.

Importantly, super-representors carry information that is not carried by $\bigcup\{[u]^{pa} : u \in U\}$. Let $U(o) = \{u(o) : u \in U\}$. Let's define the *upper* and *lower choiceworthiness* for an outcome o:

$$\text{upper choiceworthiness:} \quad \overline{U}(o) =: \max_{u \in U} u(o)$$
$$\text{lower choiceworthiness:} \quad \underline{U}(o) =: \min_{u \in U} u(o)$$

Suppose o and o' have upper and lower choiceworthinesses according to U. Then U's super-representor preserve the following form of information: for any outcomes o, o', there's some r s.t.

$$\frac{\overline{U}(o) - \overline{U}(o')}{\underline{U}(o) - \underline{U}(o')} = r$$

Notice that in the limit, there may be no precise ratios of choiceworthiness differences, or any other cardinal information, about which all elements of a choiceworthiness super-representor agree. Such a choiceworthiness super-representor can be used to represent a genuine merely ordinal moral theory. And in the other limit, if $\bigcup[U]^{pa}$ is equivalent to $[u]^{pa}$ for some CF u, then $[U]^{pa}$ contains complete precise information about ratios of choiceworthiness differences. So it counts as representing precise degrees of choiceworthiness. So the framework of imprecise choiceworthiness generalizes both merely ordinal and precise cardinal degrees of choiceworthiness in either personal preferences or moral theories. The correct decision theory for imprecise choiceworthiness, whatever it turns out to be, will be able to provide a solution to the hard problem.

We should approach this solution clear-eyed: if we represent merely ordinal theories with informationally choiceworthiness super-representors, their members will not have upper and lower bounds in their choiceworthiness assignments. (Otherwise, these would encode cardinal information.) And so only some of the imprecise metanormative decision theories I survey would be usable, and these would generally be extraordinarily permissive, prohibiting only weakly dominated options. Is this acceptable?

I assume that to say that two ordinal theories are genuinely incomparable is to say that comparisons between them are maximally indeterminate.[14] The correct imprecise metanormative decision theory should handle any degree of indeterminacy; its verdicts about merely ordinal theories should be correct. If a decision theory's verdicts in cases of uncertainty between two apparently ordinal theories are intuitively too permissive, that suggests that either

1. we aren't using the correct imprecise decision theory, or

2. the relevant moral theories aren't genuinely merely ordinal.

I suspect the latter will often be the case, for reasons noted in §2.1.

You might suspect a bait-and-switch has taken place! I've promised a resolution of the hard problem—where theories carry merely ordinal information—but it might be

[14] Some may insist on a distinction between cases where two theories are comparable, but only indeterminately so, vs. cases where two theories are incomparable in some more absolute sense. If these are distinct, there's an *even harder* problem to solve than the hard problem. Solving the harder problem is outside the scope of this paper. Thanks to [*omitted*] for discussion.

that the solution is most plausible in cases where theories carry at worst incomplete cardinal information. But this isn't quite right. Rather, the theory accommodates genuine merely ordinal theories—but it also accommodates theories that we might have confused for merely ordinal theories because their choiceworthiness assignments are incomplete. Our intuitions about decisions under uncertainty between moral theories can sometimes provide a better understanding of the theories themselves.[15]

2.3 Some decision theories for imprecise choiceworthiness

There's no consensus on the correct decision theory for imprecise credences or choiceworthiness. This is an ongoing research program. I'll mention a few candidates (framed in terms of imprecise choiceworthiness, though many of these theories are most commonly discussed for imprecise credences).[16]

Levi (1986) proposes a necessary condition on rational choices: that they be *V-admissible*. One might go further and treat this condition as also sufficient for permissibility:[17]

V-Admissibility: an act is permissible iff it is V-admissible.

> Relative to a choiceworthiness representor $U \in [U]^{pa}$ and a credence function cr, an option a is *V-admissible* just in case there's some $u \in U$ such that a's expected u-choiceworthiness relative to cr is at least as great as any alternative option a'.

V-Admissibility may be too permissive: for example, it sometimes permits rational agents to choose sequences of actions that guarantee lower choiceworthiness than some available alternative actions.[18] It can be strengthened to avoid this problem:

Caprice (imprecise choiceworthiness): A sequence of acts $\{a_1, \ldots, a_n\}$ is permissible iff there is some non-empty subset G of U such that for all $i = 1, \ldots, n$, a_i maximizes expected u for every $u \in G$.[19]

There are variants on V-admissibility that impose weaker constraints:

---

[15] Thanks to [*omitted*] for discussion.

[16] For simplicity I assume precise credences throughout.

[17] For the case of imprecise credences, Joyce (2010) defends (though does not go so far as to endorse) the analogous decision rule.

[18] See Elga (2010) for an extended discussion in the context of imprecise credences.

[19] Weatherson (manuscript) introduces and defends Caprice for imprecise credences.

Intersection maximization: an option a is permissible just in case for any alternative option a', there's some $u \in U$ such that a's expected u-choiceworthiness relative to cr is at least as great as a''s.[20]

Call the above four decision rules the *V-Admissibility Family*.

Other, significantly less permissive decision rules are available and familiar from the literature on imprecise credences:

Γ-MaxiMin *(imprecise choiceworthiness)*: for each act, determine the lowest expected choiceworthiness of performing that act across all $u \in U$, and choose an act that maximizes this quantity.

Γ-MaxiMin is a relatively pessimistic decision rule: like MaxiMin principles in general, it lets worst-case-scenarios—in this case, worst-case expectations—determine the degrees of choiceworthiness of acts. A less-discussed optimistic analogue:

Γ-MaxiMax *(imprecise choiceworthiness)*: for each act, determine the highest expected choiceworthiness of performing that act across all $u \in U$, and choose an act that maximizes this quantity.

It's also possible to introduce a class of decision rules that weigh pessimism and optimism against each other. Let $\rho \in \mathbb{R}_{[0,1]}$ be a *pessimism factor*, identifying how much weight to give U's most pessimistic assessments. Each pessimism factor determines a decision rule:

Γ-Hurwicz$_\rho$: choose an act that maximizes $\mathfrak{H}_\rho$.

$$\mathfrak{H}_\rho(x) = \rho \min_{u \in U} Eu(a) + (1-\rho) \max_{u \in U} Eu(a)$$

Notice that Γ-Hurwicz$_1$ is equivalent to Γ-MaxiMin, and Γ-Hurwicz$_0$ is equivalent to Γ-MaxiMax. Γ-Hurwicz stands to Hurwicz decision rules as Γ-MaxiMin stands to MaxiMin.

Call these last three forms of theory *the* Γ *family*.

2.4 Problems for metanormative generalizations of imprecise decision rules

I've described a formal model of choiceworthiness that can represent traditional cardinal theories, merely ordinal theories, and intermediate theories with more cardinal information than merely ordinal theories but less than precise cardinal theories. I've

[20] Intersection maximization, defended by Sen (2004), differs from V-admissibility in reversing its order of quantifiers.

also noted some candidate decision theories for imprecise choiceworthiness. So, once we settle on the correct one, have we solved the hard problem of intertheoretic comparisons?

Alas, no: there's a serious challenge to generalizing any of the decision theories above for uncertainty about moral theories.

To see why, let's consider a toy example: Suppose an agent is uncertain between t_1 and t_2, both of which are imprecise theories. She considers each equally probable. The agent has two available acts: A and B.[21] The ranges of choiceworthiness that each theory assigns these options are represented in figure 1.

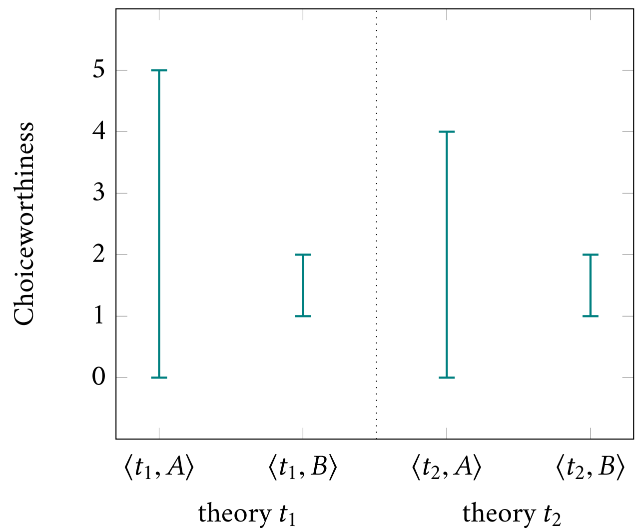

Figure 1: Interval representation of choiceworthiness

Suppose further that t_1 and t_2 can each be represented with sets containing just two CFs: t_1 can be represented with $\{u_{1x}, u_{1y}\}$ and t_2 can be represented with $\{u_{2x}, u_{2y}\}$; see figure 2.[22]

Now, suppose the agent wants to determine whether B is *intertheoretically V-admissible*. (This is an important notion: intertheoretic V-admissibility is a necessary condition on permissibility for most of the other decision rules.) *Prima facie*, in order to create an intertheoretic generalization of V-admissibility we simply need to replace traditional expected choiceworthiness with expected *intertheoretic* choiceworthiness:

[21] I use $A, B, \ldots$ as names for acts and lower-case a, with or without primes or subscripts, as variables over acts.

[22] In realistic cases, adequate representors will plausibly require infinitely many members; I use smaller representors for simplicity.

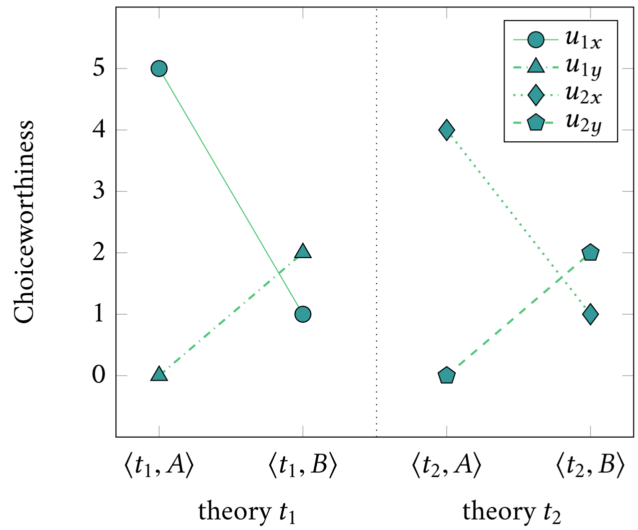

Figure 2: Imprecise degrees of choiceworthiness

Intertheoretic V-admissibility: an option is *intertheoretically V-admissible* just in case it maximizes expected intertheoretic choiceworthiness relative to some CF u in the relevant representor U.

The big question: what is the relevant representor for the intertheoretic case?

It can't merely be the union of the representors of epistemically possible theories. If it were, then as long as the option was V-admissible according to some theory that the agent had positive credence in, it would be V-admissible given the agent's normative uncertainty. So if an option maximized choiceworthiness according to any precisification of any epistemically possible theory, plausible or merely possible, then it will be V-admissible. If V-admissibility is sufficient for permissibility—probably the most widely accepted imprecise decision theory for descriptive uncertainty—such representors would lead to conflicts with expected intertheoretic choiceworthiness maximization in the special case of uncertainty between precise theories. Indeed, epistemically possible permissibility would be sufficient for genuine permissibility.

So we need a different form of representor. This will require a better understanding of the kind of CF relevant for precise expected intertheoretic choiceworthiness. In the precise case, idealizing away descriptive uncertainty, expected intertheoretic choiceworthiness was defined as follows:

$$Eiu(a) = \sum_{t \in \mathcal{T}} cr(t) u_t(a)$$

In effect, we consider two partitions of possibilities: the set of possible precise theories $\mathcal{T}$ and the set of possible worlds $\mathcal{W}$. Together, they provide a space of possibilities over which agents are uncertain: $\mathcal{T} \times \mathcal{W}$. Choiceworthiness values are then assigned to *theory-world pairs*. If we suppose every theory is representable by a precise CF, we can define an *intertheoretic choiceworthiness function* (ICF) ranging over this space of possibilities, $u_I : \mathcal{T} \times \mathcal{W} \to \mathbb{R}$: for any $t \in \mathcal{T}$ and any world w,

$$u_I(t, w) =: u_t(w)$$

And so our original definition of expected intertheoretic choiceworthiness is equivalent to a variant defined in terms of u_I:

$$Eiu_I(a) = \sum_{t \in \mathcal{T}} cr(t) u_I(t, a)$$

In the imprecise case, then, to define intertheoretic V-admissibility, we need a *set* of ICFs U_I. Then an act is intertheoretically V-admissible just in case it maximizes expected intertheoretic choiceworthiness relative to some CF u in the intertheoretic representor U_I.

A question remains: which ICFs are in this set?

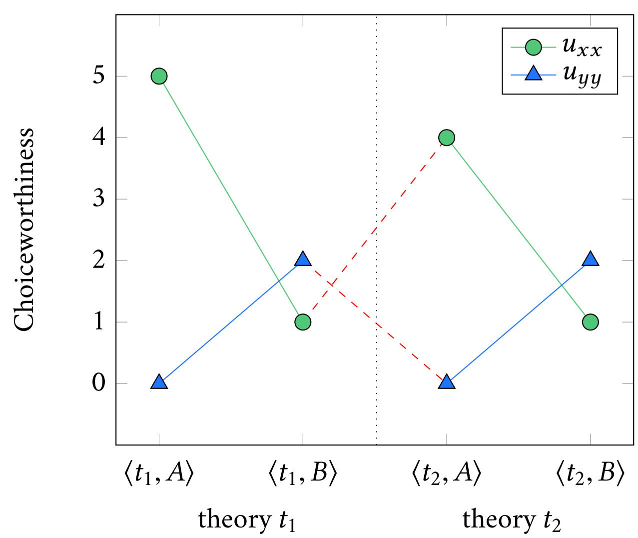

Figure 3: Two ICFs

Return to our toy example. Here's one option for how to generate a U_I from U_{t_1} and U_{t_2}. In figure 3, solid lines represent CFs; red dashed lines represent ways of combining CFs from distinct theories into unified ICFs. We define U_I as containing the following two ICFs:

$$u_{xx}(t_i, a) =: \begin{cases} u_{1x}(a) & \text{if } i = 1 \\ u_{2x}(a) & \text{if } i = 2 \end{cases}$$

$$u_{yy}(t_i, a) =: \begin{cases} u_{1y}(a) & \text{if } i = 1 \\ u_{2y}(a) & \text{if } i = 2 \end{cases}$$

The expected choiceworthiness of A and B across U_I is as follows:

	Eiu_{xx}	Eiu_{yy}
A	4.5	0
B	1	2

So both A and B are intertheoretically V-admissible for this choice of intertheoretic representor.

But consider an alternative, represented in figure 4.

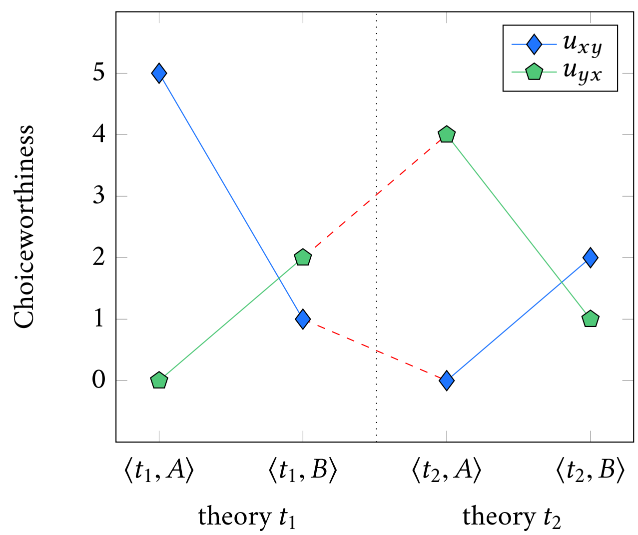

Figure 4: Alternative ICFs

Here, we define U_I^* as containing the following two ICFs:

$$u_{xy}(t_i, a) =: \begin{cases} u_{1x}(a) & \text{if } i = 1 \\ u_{2y}(a) & \text{if } i = 2 \end{cases}$$

$$u_{yx}(t_i, a) =: \begin{cases} u_{1y}(a) & \text{if } i = 1 \\ u_{2x}(a) & \text{if } i = 2 \end{cases}$$

The expected choiceworthiness of A and B across U_I^* is as follows:

	Eiu_{xx}	Eiu_{yy}
A	2.5	2
B	1.5	1.5

So for this choice of intertheoretic representor, B is not V-admissible!

What could possibly motivate any choice between $U_I = \{u_{xx}, u_{yy}\}$ and $U_I^* = \{u_{xy}, u_{yx}\}$? And what could motivate the claim that the intertheoretic representor contains two ICFs, rather than three or four?

In the case of imprecise credences, it's common to think of the elements of an agent's credal representor as members of a *committee*. Each member of the committee has precise credences, but the collective cannot faithfully be represented by any unique precise credence function: the imprecise credences are opinionated only when the members of the committee are unanimous. So similarly, if a moral theory cannot be represented with a precise CF, we might think of the theory instead as a kind of committee.

But what about intertheoretic representors, containing ICFs? It seems we're forced to take seriously the idea that, e.g., individual members of the Imprecise Utilitarian Committee correspond to individual members of the Imprecise Kantian Committee. But why would that be so? What would such correspondences below the level of the representor model in moral reality?

In effect, we represented the theories' committees as having members in common. But why would that be so? These theories represent *competing* ways of viewing the world. The moral theories at issue are mutually exclusive *by design*; expectations require it. Prima facie, their committees aren't compatible, and shouldn't overlap. Committee members can't serve on more than one. These aren't like departmental committees, cooperating in a division of labor. They're instead like US political parties: they are competitors; they cannot both win.

But notice that this problem is *not* merely a version of the easy problem of intertheoretic comparisons. Solving the easy problem required finding an appropriate shared *scale*, such that it would make sense to make comparative claims like $u_{t_i}(A) = u_{t_j}(B) - 5$. But in the imprecise case, even if we can compare the choiceworthiness assignments of any two members of any competing theories' representors, that does nothing to address the present problem. Even if there's a fact of the matter about how, e.g., the upper and lower choiceworthinesses of a according to t_1 and t_2 compare, that tells us nothing about how the two theories are related at the sub-representor level, in their internal structure: how specific elements of U_{t_1} correspond to specific elements of U_{t_2}.

Here's another way to describe the problem: imprecision represents ways in which a theory's "preferences" are indeterminate or incomplete. When we represent such a theory with a set of CFs, those functions are meant to be something like ways in which the theory's assignments of choiceworthiness could be *completed*, or sharpenings that are in some sense compatible with the theory. But it's unclear how, or whether, possible completions of one theory's choiceworthiness assignments place any constraints on the possible completions of another theory's choiceworthiness assignments. Prima facie, however t_1 might be sharpened has nothing to do with how t_2 might be sharpened.

In short: the worry is that generalizing decision theories under discussion for moral uncertainty would require reifying elements of the model that *don't correspond to anything in moral reality*.

I consider three responses to this problem in the next sections.

3 Imprecise metanormative decision theory: some options

3.1 Unstructured rules

We can avoid the puzzle raised in §2.4 by using decision rules that don't require intertheoretic representors. A natural strategy is to construct decision rules that are sensitive, not to the specific set of choiceworthiness assignments to an outcome, but instead to the *upper* and *lower bounds* of these assignments. These rules, which I'll call *unstructured rules,* ignore the interior structure in choiceworthiness representors. They ignore intermediate choiceworthiness assignments between the upper and lower bounds, or relations among sharpenings for different outcomes. They revert back to only information carried in the interval representations of theories. (Contrast these with *structured rules,* like the V-Admissibility family and the Γ family, that look to the interior structure of imprecise choiceworthiness assignments.)

First, let's distinguish the *lower expected choiceworthiness* of an act from the *expected lower choiceworthiness* (expected $\underline{U}$) of an act. The first notion uses a set of ICFs and a credence function to generate a set of expectations of choiceworthiness, and then selects the minimum of these expectations. The second notion looks at the lower choiceworthiness for each possible outcome of an act (where these infima might be assigned by different precise CFs), and then takes the expectation of these values. Formally:

Lower expected intertheoretic choiceworthiness: $\underline{E}u(a) = \min_{u_I \in U_I} \sum_{t \in T} cr(t) u_I(t, a)$

Expected lower intertheoretic choiceworthiness: $E\underline{U}(a) = \sum_{t \in T} cr(t) \min_{u \in U_t} u(a)$

We can characterize some candidate decision rules. The first is a variation on Γ-MaxiMin:

Ξ-**MaxiMin:** choose an act that maximizes expected lower intertheoretic choiceworthiness.[23]

Recall that Γ-MaxiMin tells agents to maximize lower expected intertheoretic choiceworthiness. To see that Ξ-MaxiMin and Γ-MaxiMin are not equivalent, consider an example. In table 1, we consider two acts, two theories, and an intertheoretic representor containing two ICFs. Suppose the theories are equiprobable.

		t_3	t_4	expected $u_i(a)$ for Γ-MaxiMin	expected $\underline{U}(a)$ for Ξ-MaxiMin
A	u_1	14	2	8	**3**
	u_2	4	6	5	$(\underline{U}(A, t_3) = 4, \underline{U}(A, t_4) = 2)$
B	u_1	8	4	**6**	2
	u_2	0	12	**6**	$(\underline{U}(B, t_3) = 0, \underline{U}(B, t_4) = 4)$

Table 1: Γ-MaxiMin vs. Ξ-MaxiMin

Γ-MaxiMin recommends B, while Ξ-MaxiMin recommends A.

Ξ-MaxiMin is a pessimistic principle: everything is decided by the lower bounds of the epistemically possible choiceworthiness for outcomes. There is a corresponding optimistic principle, which is similarly related to Γ-MaxiMax. Where expected upper intertheoretic choiceworthiness is defined on the model of expected lower intertheoretic choiceworthiness,

Ξ-**MaxiMax:** choose an act that maximizes expected upper choiceworthiness.

Finally, there is a variation on Γ-Hurwicz for upper and lower expectations:

Ξ-**Hurwicz**$_\rho$**:** for pessimism factor $\rho \in \mathbb{R}_{[0,1]}$, choose an act that maximizes $\mathfrak{H}_\rho$.

Intertheoretic $\mathfrak{H}_\rho(a) =: \rho \cdot E\underline{U}(a) + (1 - \rho) \cdot E\overline{U}(a)$

Call these unstructured decision rules *the* Ξ *family.* These avoid the reification worries in §2.4: they do not require the assumption that any sharpenings of the representor for one theory must correspond to any particular sharpenings of the representor for a different theory.

The Ξ family has questionable results, though. These rules are incapable of distinguishing between the following kinds of cases:

[23] Gilboa & Schmeidler (1993) defend this decision rule for imprecise credences.

- cases where one theory determinately assigns higher choiceworthiness to an outcome than does another theory, even though they assign overlapping ranges of choiceworthiness—e.g., t_3's vs. t_4's assessment of A in figure 5.
- cases where there is no determinate comparison between two theories' choiceworthiness assignments to an outcome.—e.g., t_3's vs. t_5's assessment of A in figure 5.

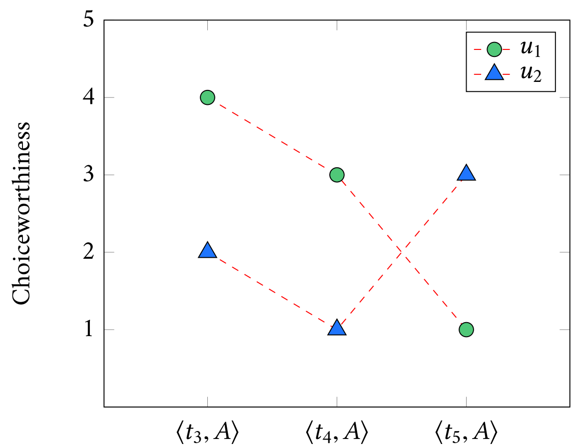

Figure 5: Determinate vs indeterminate intertheoretic comparisons

There are, plausibly, scenarios where two theories t and t' have enough in common that it can be determinate that t regards an outcome a more highly than t' does (even when their choiceworthiness assignments are imprecise and overlapping). We'll say that, in such cases, t is *more sanguine* than t' about a. There are a variety of ways in which two theories might compare in this way: for example, t might be an *amplified* version of t'.[24]

There's a reason why the analogous rules aren't popular in the case of imprecise *descriptive* uncertainty. Consider an analogous descriptive case. Suppose you assign imprecise degrees of choiceworthiness, in overlapping ways, to going to the Taiwanese restaurant (T) and going to the Indian restaurant (I). You're not merely indifferent between the two: if you were offered a \$1 discount for the Taiwanese restaurant (T^+), you would continue to be conflicted about the choice between the two restaurants.[25]

[24] For a defense of amplified theories and examples, see MacAskill et al. (2020), 125–130.

[25] For discussion of such cases of "insensitivity to sweetening," see Chang (1997); Hare (2010); Schoenfield (2014); Doody (2019a,b).

Yet you determinately prefer going to the Taiwanese restaurant with the discount over going there without the discount. Your choiceworthiness assignment are as follows: for some small positive ε,

		u_1	u_2
(T^+)	Taiwanese+$1	$3+\varepsilon$	$1+\varepsilon$
(T)	Taiwanese	3	1
(I)	Indian	1	3

Unstructured rules, which attend only to upper and lower choiceworthiness for each option, face a problem with this sort of case: given that (T) and (I) have the same upper and lower choiceworthiness, the rules cannot accommodate their different relations to (T^+): (T^+) is determinately preferable to (T) but *not* determinately preferable to (I).

The Ξ Family faces other concerns: in particular, that Ξ-MaxiMin (-MaxiMax) requires choices that are extreme in their risk-avoidance (-inclination), even in comparison with the corresponding Γ rules—particularly once we include descriptive uncertainty. The range of possible outcomes for an act have greater or equal variance in choiceworthiness to the range of expectations for the act. So Ξ-MaxiMin's recommendations are based on averages of more pessimistic quantities than Γ-MaxiMin's (as table 1 illustrates). Mutatis mutandis for Ξ-MaxiMax.

3.2 Preserve structured rules with Maximalism

3.2.1 Maximalism

So much for unstructured rules. There's an alternative response to the puzzle raised in §2.4: we can insist that there must exist an intertheoretic representor—reification worries be damned—but aim to avoid arbitrary choices about which ICFs it includes. The risk of arbitrariness is minimized if we impose a *Maximalism* constraint on the intertheoretic representor: within the intertheoretic representor, every element of every representor for every theory forms a unified ICF with every element of every other choiceworthiness representor for every other theory.[26] In the example from §2.4, this would mean the intertheoretic representor would contain all four of the ICFs, as in figure 6.

The primary benefit of using intertheoretic representors is that it allows us to generalize any of the traditional decision rules for imprecise choiceworthiness discussed in §2.3. These generalizations all make use of the notion of individual members of an

[26] More precisely, for a set of theories $\mathcal{T} = \{t_1, \ldots, t_n\}$, each t_i has an informationally adequate U^{t_i}. These determine a set of sequences $U^{t_1} \times \ldots \times U^{t_n}$. For each $\mathbf{u}$ in this set, the intertheoretic representor U_I contains an ICF u_I such that for all i, $1 \leq i \leq n$, and all $w \in \mathcal{W}$, $u_I(w, t_i) = \mathbf{u}_i(w)$.

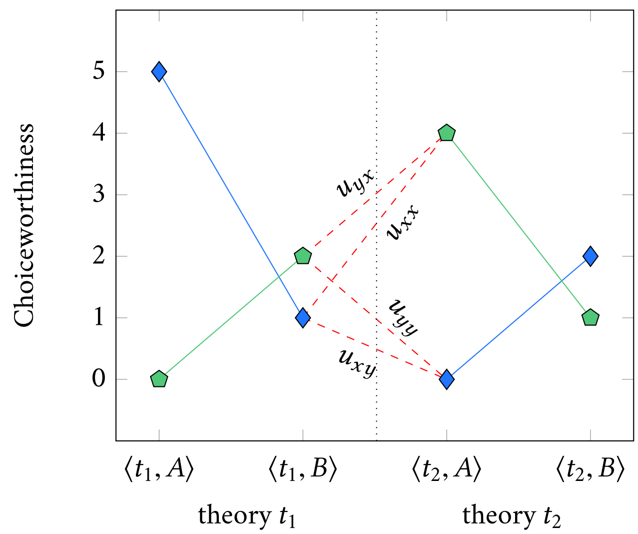

Figure 6: Intertheoretic representors under Maximalism

intertheoretic representor. Call these *structured rules*: they take into consideration the interior structure of intertheoretic choiceworthiness representors.

But Maximalism is not a wholly normatively innocent assumption.

3.2.2 Problems for the Maximalist constraint

Once we impose the Maximalism constraint on intertheoretic representors, our structured rules—like the unstructured rules—are insensitive to some kinds of differences in how a pair of theories might be interrelated. For example, they'll be insensitive to the possibility that one theory is determinately more sanguine than another about an outcome when their choiceworthiness assignments are overlapping. Maximalism rules out this possibility.

Let's suppose two theories, t_1 and t_2, both agree entirely about the grounds and nature of self-realization's contribution to choiceworthiness, so that when two options only exhibit self-realization as a source of contributory value, it's appropriate to treat t_1 and t_2 as assigning the same imprecise choiceworthiness to that option. Suppose also that t_1 also sees happiness as contributing to choiceworthiness in a way that t_2 doesn't: more happiness is better. The contributions of self-realization and happiness to choiceworthiness are separable: an increase or decrease in one source of value never affects the other.

Suppose an outcome A produces some degree of both self-realization and happiness. So t_1 sees it as having all the self-realization value as t_2, but as receiving some extra contribution of hedonistic value to its overall degree of choiceworthiness. t_1 and t_2's choiceworthiness assignments to A are imprecise and overlapping. Nevertheless, t_1 is determinately more sanguine about A than t_2 is.

So it ought to be the case that all ICFs in the relevant intertheoretic representor assign a higher degree of choiceworthiness to A in the event that t_1 is true than in the event that t_2 is true. In other words, the relevant representor ought to contain, say, only u_1 and u_2 in figure 7. But the Maximalism constraint holds that u_3 must be included. This amounts to insisting that it's *indeterminate* whether t_1 is more sanguine about A than t_2 is.

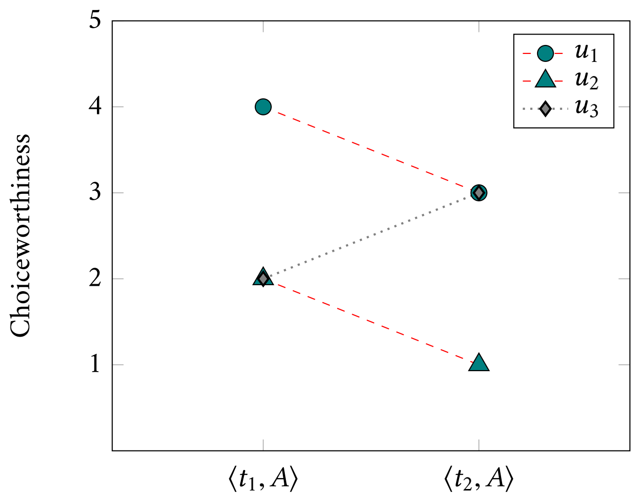

Figure 7: Determinate comparative sanguinity

Notice that here, we are a considering a case where, intuitively, sharpening one theory *does* place constraints on how another is sharpened: however we sharpen t_1, we need to sharpen t_2 such that its choiceworthiness assignment to A is lower. It's unclear how widespread such cases may be. They seem to arise only for theories that have substantive agreement about the grounds and nature of at least some choiceworthiness assignments. It's a problem for Maximalism, qua model of intertheoretic comparison, that it fails to accommodate this form of determinate comparative sanguinity.

Both the V-Admissibility family and the Γ family of decision rules face problems when paired with Maximalism. Consider the following example: Suppose your confidence is divided evenly between t_3 and t_4. t_3 holds that overall happiness and generosity are both sources of contributory value, while t_4 holds that overall happiness and fiscal responsibility are both sources of contributory value. On both views, happiness contributes a determinate amount of choiceworthiness to an option, while their respective other source's contribution is indeterminate. You're considering an option A—say, buying a small gift for a friend—that increases overall happiness somewhat, and also exhibits both some measure of fiscal responsibility and some measure of generosity. On both views, A's degree of choiceworthiness is indeterminate.

Without the Maximalism constraint, we might represent both theories with a pair of ICFs, u_1 and u_2. u_1 is sensitive to happiness and, if it is valuable, fiscal responsibility; u_2 is sensitive to happiness and, if it is valuable, generosity. u_1 regards A highly if t_4 is true: if fiscal responsibility increases choiceworthiness. u_2 regards A highly if t_3 is true: if generosity is valuable. So their comparative assessments of A across theories are reversed.

Suppose you're choosing between A and B, where B derives no contributory value from generosity or fiscal responsibility, but generates slightly more happiness than A. Figure 8 represents u_1 and u_2 across these two options.

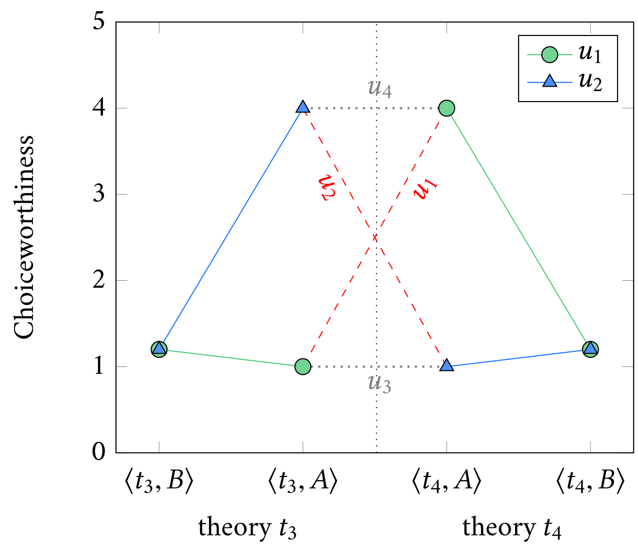

Figure 8: Maximalist intertheoretic representor

Now, without the Maximalism constraint, both V-Admissibility and Γ-Maximin require preferring A to B: on both u_1 and u_2, the expected choiceworthiness of A is greater than that of B. But Maximalism requires that we include in the intertheoretic representor another pair of CFs: u_3 and u_4, as represented in figure 8. With these additions, A is no longer required. V-Admissibility now declares B permissible (since it maximizes expected choiceworthiness according to u_3). And Γ-MaxiMin declares B obligatory (since the introduction of u_3 gives A a *lower* lowest expected choiceworthiness).

In this case, this effectively treats *both* generosity and frugality as *valueless*—an inaccurate representation of your decision situation, where you're certain that at least one of the two affects choiceworthiness, albeit indeterminately. (Analogous examples may be constructed for Γ-MaxiMax and Γ-Hurwicz rules.)

3.3 Structure without Maximalism

At this stage, we might conclude that intertheoretic representors are an acceptable modeling tool, *pace* the reification worry, and that the maximalism constraint on intertheoretic representors is false. Sometimes, a member of one theory's representor forms unified ICFs with *some*, but not *all*, members of another theory's representor.

This is, by my lights, a fine view to land on! But it's worth considering a modification of this view, generalizing Carr's (2020) resolution to the easy problem. This alternative will be able to make the same kinds of predictions as the structured rules, but won't invite the same kind of reifying concerns for intertheoretic representors.

To motivate this proposal, we'll consider some puzzles for non-maximalist intertheoretic representors.

3.3.1 Epistemic problems

Recall our example represented in figure 7. An agent may be certain that t_1 and t_2 are related in such a way that their intertheoretic representor excludes, say, u_3: certain that t_1 is determinately more sanguine than t_2 about A. An agent may instead be certain that t_1 and t_2 are related in such a way that u_3 is included: certain that neither theory is determinately more sanguine than the other about A.

But what about an agent who isn't certain how the theories relate? If we are willing to countenance rational moral uncertainty, we may well also accept that rational agents need not know how theories relate below the level of representors. Agents may be uncertain, for example, of whether one imprecise theory is determinately more sanguine about an outcome than another.

Earlier, our worry was *metaphysical*: that such connections don't exist. Now, our worry is *epistemic*: it's not clear whether or how our uncertainty about which ICFs are in the intertheoretic representor should be modeled, and whether or how it should affect which actions are appropriate.

We might simply include every ICF the agent cannot rule out. But the intertheoretic representor isn't meant to represent the agent's state of uncertainty. The agent's state of uncertainty is represented in her credences. So this would generate a redundancy in the model.

Worse, the point of metanormative decision theories is to represent rational behavior under uncertainty about something distinctly non-subjective: the *moral truth*.

The form of ICF used in the precise case represents objective commensurations of distinct theories of the objective moral truth. In the imprecise case, the intertheoretic representor is meant to do the same. We cannot justify including ICFs in the objective intertheoretic representor on the grounds that the agent is subjectively, epistemically unable to rule them out.

This complaint sounds fussy, but it has serious consequences: which ICFs are included can have radical effects on which actions are morally appropriate. The inclusion of one extra ICF can change an action's deontic status from impermissible to permissible.

3.3.2 *De dicto* imprecise choiceworthiness

The problem in §3.3.1 arises only under a specific assumption: that the form of normative uncertainty relevant for metanormative decision rules is uncertainty between full-fledged moral *theories*. Once we represent moral theories with choiceworthiness representors, questions about how these representors compare, and how their members are related, become inevitable. Instead, I propose, we can proceed with *local*, incomplete, and comparative moral hypotheses.

To accomplish this, we can use a generalization of Carr's (2020) *de dicto* expectations of choiceworthiness. Instead of competing CFs, Carr uses one CF specified under a description: "the genuine moral CF u" (whatever that is). The relevant state space isn't a set of theories, but a set of hypotheses of the following form: for an act a and real number r, *the genuine degree of moral choiceworthiness of a is r*.

Crucially, with *de dicto* choiceworthiness, the possible values of outcomes are not defined in relation to *theories*, but instead defined in relation to each other: *'aing is okay, though nothing to applaud'* vs. *'aing is not okay, though not evil'* vs. *'aing is downright monstrous'*…

The choiceworthiness relations between these hypotheses isn't something to *discover*, as with complete moral theories, but instead something to *stipulate*. These stipulations have a reference-fixing function: they determine which possibility is under discussion. (*'The possibility that the genuine degree of choiceworthiness of a is r'* vs. '… $r-1$' vs. …)

With localist, *de dicto* choiceworthinesses, the form of uncertainty described in §3.3.1, uncertainty about sub-representor relations between theories, isn't modeled any differently from uncertainty between theories themselves: *'aing is indeterminately choiceworthy in* this *way'* vs. *'aing is indeterminate but definitely worse'* vs. *'aing is indeterminate in an incomparable way'*… So §3.3.1's epistemic worry doesn't arise on this model.[27]

[27] *Objection:* Can't the proponent of intertheoretic representors also use this maneuver, insisting that the

3.3.3 Formalization

In this section, I'll briefly offer a more general model of imprecise *de dicto* choiceworthiness and how we can extract notions like V-admissibility from it. This model generalizes that in Carr (2020).

Some notation: let $\mathbb{U}$ represent the set of all CFs. $\mathbb{U}^n$ is the set of all ordered n-tuples of CFs. The relevant n is determined by our modeling needs for capturing the agent's state of uncertainty over imprecise theories.[28] For the agent with imprecise moral uncertainty, there's some subset S of $\mathbb{U}^n$ such that each $\mathbf{U} \in S$ might, for all she knows, contain all and only CFs in *the genuine moral representor*.[29] Adapting Carr (2020), our notation for the genuine moral representor uses typewriter font: '$\mathtt{U}$'. Note that '$\mathtt{U}$' is not a name but a definite description: the genuine choiceworthiness representor specified *de dicto*.

Using ordered n-tuples will allow us to model the kinds of uncertainty over imprecise choiceworthiness at issue without reifying ICFs. For example, suppose we model an agent as uncertain between a pair of candidates for $\mathtt{U}$, $\mathbf{U}$ and $\mathbf{U}'$, s.t. for all i, $1 \leq i \leq n$, the ith item in $\mathbf{U}'$ assigns lower precise choiceworthiness to an outcome o than the ith item in $\mathbf{U}$ (notation: $u'_i(o) < u_i(o)$). This doesn't commit us to a claim about the relationship between two independently characterized moral theories. It's simply a way of modeling a state of uncertainty for an agent who's uncertain whether o's degree of choiceworthiness is mushy in one way, or in another way that's determinately worse.

Each option a will have an ordered n-tuple of $E\mathtt{u}_1(a), \ldots, E\mathtt{u}_n(a)$ of expected *de dicto* choiceworthinesses. Each $E\mathtt{u}_i$ corresponds to the ith coordinates of the $\mathbf{U} \in \mathbb{U}^n$. In the limiting case, where choiceworthiness is precise, $n = 1$ and this is equivalent to precise expected *de dicto* choiceworthiness.[30]

---

relevant form of uncertainty only arises if we haven't individuated theories finely enough? *Reply:* This would require decision-theoretically significant distinctions between theories with completely identical choiceworthiness representors.

[28] I assume finitude for simplicity. The relevant n might be determined by the products of the cardinalities of all epistemically possible imprecise choiceworthiness assignments to all epistemically possible outcomes—which may be unwieldy. Because the view presented has a localist spirit, we may not need so great an n in specific contexts to capture the believer's state of uncertainty.

[29] I assume that the relevant scale for selecting members of super-representors has been fixed. For details on how this is accomplished, see (Carr, 2020), §3.2 and §4.1.

[30] If this looks too similar to intertheoretic uncertainty, we can construct a more localist model. Each hypothesis about the distribution of imprecise choiceworthinesses over options is represented with a matrix. Suppose there are m possible acts in the relevant choice situation with n (possibly redundant) choiceworthiness sharpenings. (I assume purely normative uncertainty for simplicity.) We'll represent a local hypothesis about imprecise choiceworthinesses with a matrix $\mathbf{V} \in \mathbb{R}^{m \times n}$, where the ith row represents a_i's range of imprecise choiceworthiness.

$$Eu_i(a) = \sum_{\substack{s \in \mathcal{S} \\ \mathtt{U}=\langle u_1,\ldots,u_n\rangle \in \mathbb{U}^n}} cr(s \wedge \mathtt{U} = \mathbf{U} \mid\mid a) u_i(s \wedge a)$$

We end up with a sequence of expected choiceworthinesses. We can define a family of *de dicto* decision rules:

V-Admissibility *de dicto*: an option is permissible iff it's V-admissible *de dicto*.

An option a is *V-admissible de dicto* iff there's some i such that for all alternative options a', $Eu_i(a) \geq Eu_i(a')$.

Intersection Maximization *de dicto*: an option a is permissible iff for all alternative options a', there's some i such that $Eu_i(a) \geq Eu_i(a')$.

Γ-**Maximin *de dicto*:** an option a is permissible iff $\min\{Eu_i(a) : 1 \leq i \leq n\} \geq \min\{Eu_i(a') : 1 \leq i \leq n\}$ for all alternative options a'.

And so on.

The improvement here is conceptual rather than technical. Those who are comfortable with general intertheoretic representors, and unconcerned about either the metaphysical or epistemic worries I raised, can make equal use of this formalism.

The conceptual improvement comes with the possibility of constructing a spread of metanormative hypotheses that are determined *locally, in relation to each other,* rather

---

$$\mathbf{V} = \begin{pmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,n} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m,1} & v_{m,2} & \cdots & v_{m,n} \end{pmatrix}$$

Suppose an agent is uncertain a_1's choiceworthiness, but certain that a_1 is determinately more choiceworthy than a_2. Then all epistemically possible **V**'s i's entry for a_1, v_1, i, will be greater than its entry for a_2. If the agent is uncertain whether a_1 is determinately more choiceworthy than a_2, then some epistemically **V** won't have that property.

Each **V** represents a *conjunction* of hypotheses about the choiceworthiness of different options. Where $v_{i,*}$ is the i^{th} row of **V**, we can let $H(\mathbf{V})$ be the proposition: $\mathtt{U}(a_1) = v_{1,*} \wedge \ldots \wedge \mathtt{U}(a_n) = v_{n,*}$.

Then we can define a sequence of expected *de dicto* choiceworthinesses as follows:

$$Eu_j(a_i) = \sum_{\mathbf{V} \in \mathbb{R}^{n \times m}} cr(H(\mathbf{V})) \cdot v_{i,j}$$

For each epistemically possible **V**, $Eu_j(a_i)$ looks at **V**'s imprecise choiceworthiness for option a_i and picks out the jth sharpening. $Eu_j(a_i)$ then takes a weighted average of the possible jth sharpenings, weighted by the probability of each hypothesis **V** about the distribution of choiceworthinesses across options.

than in relation to theories. The open questions are local questions about how the choiceworthiness of different options might compare with each other. There's no question of sub-representor intertheoretic unifications here. There's only a question of how confident the agent is in different hypotheses about how options relate to each other: whether one might or might not be determinately better than the other, to what degrees, in what ways.

3.4 *A bonus: precise theories that are imprecisely comparable*

Using imprecise choiceworthiness assignments, *de dicto* or otherwise, has an added bonus: it allows for modeling the relation between *precise* theories that are *imprecisely* comparable, and constructing decision theories for uncertainty over these.

A toy example: suppose it's indeterminate whether t_1's assignments to A and B is greater than t_2's, even though their preferences regarding these two outcomes are reversed. These theories are related as in figure 9. But despite being represented with two choiceworthiness functions, neither theory is imprecise. Rather, each theory is redundantly represented with two informationally equivalent CFs. But since the relation between the two theories is indeterminate, there is one way of sharpening their relation such that, e.g., t_1 assigns A higher choiceworthiness than does t_2, and another where the reverse is true. So we represent the imprecise relation between precise theories using an intertheoretic representor, containing ICFs representing both sharpenings. Then we apply our choice of imprecise metanormative decision theory.

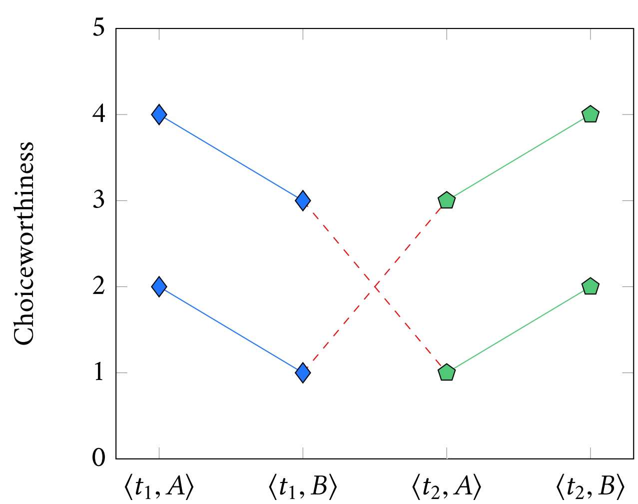

Figure 9: Imprecise relations between precise theories

4 Conclusion

The hard problem of intertheoretic comparisons requires a metanormative decision theory for moral uncertainty between cardinal and merely ordinal theories. I've argued that imprecise choiceworthinesses can accommodate both, as well as a range of intermediate forms of theory. But in order to generalize imprecise decision theories for moral uncertainty, we face a puzzle: these decision theories put great weight on parts of our decision theoretic models that don't obviously represent anything in moral reality. I've offered a variety of options for how to handle the puzzle.

References

Carr, Jennifer Rose (2020). "Normative Uncertainty Without Theories." *Australasian Journal of Philosophy* 98(4), 747–762.

Chang, Ruth (1997). *Incommensurability, Incomparability and Practical Reason*. Harvard University Press.

Doody, Ryan (2019a). "Opaque Sweetening and Transitivity." *Australasian Journal of Philosophy* 97(3), 559–571.

Doody, Ryan (2019b). "Parity, Prospects, and Predominance." *Philosophical Studies* 176(4), 1077–1095.

Elga, Adam (2010). "Subjective Probabilities Should Be Sharp." *Philosophers' Imprint* 10(05),

Gilboa, Itzhak & David Schmeidler (1993). "Updating Ambiguous Beliefs." *Journal of Economic Theory* 59. 33–49.

Hare, Caspar (2010). "Take the Sugar." *Analysis* 70(2), 237–247. doi:10.1093/analys/anp174.

Harman, Elizabeth (2011). "Does Moral Ignorance Exculpate?" *Ratio* 24(4), 443–468.

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

Joyce, James M. (2010). "A Defense of Imprecise Credences in Inference and Decision Making." *Philosophical Perspectives* 24(1), 281–323.

Levi, Isaac (1986). *Hard Choices: Decision Making under Unresolved Conflict*. New York: Cambridge University Press.

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

MacAskill, William (2014). *Normative Uncertainty*: University of Oxford dissertation. https://www.academia.edu/8473546/Normative_Uncertainty.

MacAskill, William (2016). "Normative Uncertainty as a Voting Problem." *Mind* 125(500), 967–1004.

MacAskill, William, Krister Bykvist & Toby Ord (2020). *Moral Uncertainty*. Oxford: Oxford University Press.

Martinez, Joseph (Unpublished). "The Borda Rule and Normative Uncertainty."

Oddie, Graham (1994). "Moral Uncertainty and Human Embryo Experimentation." In K. W. M. Fulford, Grant Gillett & Janet Martin Soskice (eds.), *Medicine and Moral Reasoning*, 3–144. Cambridge University Press.

Phillips-Brown, Milo (forthcoming). "What Does Decision Theory Have to Do with Wanting?" *Mind* .

Ross, Jacob (2006). "Rejecting Ethical Deflationism." *Ethics* 116(4), 742–768.

Schoenfield, Miriam (2014). "Decision Making in the Face of Parity." *Philosophical Perspectives* 28(1), 263–277. doi:10.1111/phpe.12044.

Sen, Amartya (2004). "Incompleteness and Reasoned Choice." *Synthese* 140(1-2), 43–59. doi:10.1023/B:SYNT.0000029940.51537.b3.

Sepielli, Andrew (2009). "What to Do When You Don't Know What to Do." *Oxford Studies in Metaethics* 4. 5–28.

Sepielli, Andrew (2013). "Moral Uncertainty and the Principle of Equity Among Moral Theories." *Philosophy and Phenomenological Research* 86(3), 580–589.

Tarsney, Christian (2018). "Intertheoretic Value Comparison: a Modest Proposal." *Journal of Moral Philosophy* 15. 324–344.

Weatherson, Brian (manuscript). "Decision Making with Imprecise Probabilities."