

The Value of Normative Information*

Jeffrey Sanford Russell

July 2021

Abstract

This paper explores the idea that it is instrumentally valuable to learn normative truths. We consider an argument for “normative hedging” based on this principle, and examine the structure of decision-making under moral uncertainty that arises from it. But it also turns out that the value of normative information is inconsistent with the principle that learning *empirical* truths is instrumentally valuable. We conclude with a brief comment on “metanormative regress.”

Learning normative decision theory seems like it would be practically valuable. It is reasonable to expect making better decisions to turn out better, on balance, than making worse decisions. But we have a lot of uncertainty about which decisions are better than others, and some of this is attributable to normative uncertainty. Is risk-aversion rational? Should we assign any outcomes infinite utility? What about the St. Petersburg paradox, or Newcomb’s puzzle, or Death in Damascus? We expect that improving our understanding of normative issues will help guide us to make better decisions. And we expect that it is worth spending some effort to improve our understanding of decision theory—that improving our epistemic situation with respect to normative questions is instrumentally valuable.

Is this right? It turns out that if it is, it has striking implications for *decision-making under normative uncertainty*. First (section 1), the principle that normative information is instrumentally valuable provides new motivation for

*Thanks to the participants in an April 2020 (virtual) seminar in epistemology at USC for their very helpful feedback (Andrew Bacon, Simon Blesseohl, Irene Bosco, Stephanie Van Fossen, Jeremy Goodman, Zach Goodsell, Laura Gurskey, John Hawthorne, Yoaav Isaacs, Vishnu Sridharan, and Levy Wang). Special thanks to Simon Blesseohl, Yoaav Isaacs, Brian Jabarian, Christian Tarsney, and an anonymous referee for their comments.

the idea that decision theory should give normative uncertainty a special role: that one should “hedge” between alternative normative theories when one is uncertain which one is true. It also turns out (in section 2) to be inconsistent with straightforward applications of the most popular rules that have previously been offered for handling normative uncertainty; but it guides us toward a way of refining these rules, and toward structural constraints a “metanormative” theory must satisfy. In section 3 we consider a stronger principle in the same spirit that imposes tighter constraints on such theories. But in section 4 this strategy for motivating a theory of normative uncertainty runs into a problem: preferring to know the answers to normative questions is, in general, inconsistent with also preferring to know the answers to empirical questions. In section 5 we also briefly consider implications of the value of normative information for uncertainty about “metanormative” principles.

I will focus on decision-theoretic uncertainty (for discussion see Nozick 1993, 43–50; MacAskill 2016)—indeed, for the most part, on uncertainty about a single question in decision theory: whether risk-aversion is rational (see also Dietrich and Jabarian, manuscript, 7 and sec. 4; manuscript). This narrow focus will help keep things tractable, but it is meant to illuminate broader points about uncertainty about normative matters, including *moral* uncertainty—such as uncertainty about how good or bad things are intrinsically, or about whether it is especially wrong to violate rights, or the moral worth of animals, and so on.¹ These are questions we would like to know the answers to, not merely out of intellectual interest, but out of *ethical* interest. The choice to try to learn the answers to such questions is itself morally important (see MacAskill, Bykvist, and Ord 2020, ch. 9). It is surprisingly challenging to develop a theory that rationalizes it.

1 Don’t tell me what to do

Let’s start with a simple example (table 1). Rita faces a choice. If she chooses to *Gamble*, she gives up a sticker, and if a fair coin comes up heads she wins a free coffee mug. If she chooses to *Pass* she just keeps her sticker either way. She assigns utility 96 to winning the mug, utility 32 to just keeping her

¹See Hudson (1989); Lockhart (2000); Ross (2006); Guerrero (2007); Sepielli (2009); MacAskill (2014a), especially ch. 7; MacAskill (2014b); MacAskill and Ord (forthcoming); Podgorski (2020); MacAskill, Bykvist, and Ord (2020); for overview see Bykvist (2017); for critical discussion see Harman (2015); Weatherson (2014); Weatherson (2019); Hedden (2016)

sticker, and utility 0 to losing out on both. Should she take the gamble?²

Table 1: Utilities for Rita’s basic gamble.

	Heads	Tails
Gamble	96	0
Pass	32	32

Standard expected utility theory (EU) clearly says that Rita should Gamble. But standard expected utility theory is not the only game in town. Some decision theorists argue that it can be rational to be *risk-averse*.³ If this is right, then it can be permissible to Pass. For concreteness, suppose that the only alternative to standard risk-neutral expected utility theory is *risk-weighted expected utility theory*, where Rita’s “risk function” is, in particular, $r(p) = p^2$. Call this REU(p^2).⁴ According to this theory, Gamble’s value is 24, which is less than the value of Pass, 32.⁵

(To keep things concrete, in this example I am supposing that Rita assigns numerical utilities on a single scale for both decision theories. These assumptions will be dispensed with later.)

Rita knows that one of these two decision theories is correct—EU or REU(p^2)—but she doesn’t know which.⁶ Rita considers each theory equally likely, with probability 1/2.

²This is a gamble at 2 : 1 odds, scaled up by 2⁵ to avoid fractions in subsequent calculations.

³See for example Allais (1953); Quiggin (1982); Quiggin (1993); Machina (1989); Buchak (2013). Throughout, by “risk-averse” I mean risk-averse with respect to *utility*. Standard EU does permit risk-aversion with respect to goods like money or longevity, but not with respect to utility.

⁴Buchak (2013); economists know this as *rank-dependent utility theory* (RDU) (Quiggin 1993).

⁵Here’s how to calculate the REU (RDU) of an act. List the possible utilities in ascending order: $u_0 < u_1 < \dots$. For $i \geq 1$, let p_i be the cumulative probability of attaining *at least* utility u_i . Then the REU is

$$u_0 + r(p_1) \cdot (u_1 - u_0) + r(p_2) \cdot (u_2 - u_1) + \dots$$

In this simple case,

$$\text{REU}(\text{Gamble}) = 0 + r(1/2) \cdot (96 - 0) = (1/2)^2 \cdot 96 = 24$$

⁶If the correct decision theory is true a priori, then Rita falls short of “ideal rationality,” in

Rita has a tough decision to make. But she is in luck: she has the opportunity to consult a decision theory oracle. If she asks, the oracle will tell her which of the two decision theories is correct, and after that she can choose whether to take the gamble. This seems like a good opportunity: if she doesn't consult the oracle, she doesn't know whether she will choose rationally, but if she *does* consult the oracle, she can guarantee that she will make the best choice. What's not to like?

To be clear: we are currently interested in the *instrumental* value of normative information. Even if Rita enjoys knowing truths about decision theory or acting rationally for their own sake, we are setting such considerations aside.

The standard tool for analyzing this kind of question is the *value of information*.⁷ We can consider "consult the oracle and then decide" as a complicated act of its own. Rita knows that if she learns that EU is correct, she will choose Gamble (since she knows this maximizes expected utility), and she knows that if she learns that REU is correct, she will choose Pass (since she knows this maximizes risk-weighted expected utility). So her "oracle act" will have the same result as Gamble if EU is correct, and it will have the same result as Pass if REU is correct.

So Rita now has three options: Pass, Gamble, or Oracle. The pay-offs of these options are in table 2. Each state has probability 1/4. (Which decision theory is correct is independent of the outcome of the coin flip, on Rita's evidence.)

Table 2: Utilities for Rita's three options

	EU, Heads	EU, Tails	REU, Heads	REU, Tails
Gamble	96	0	96	0

that there are a priori truths she does not know. But this cognitive limitation does not make serious trouble. It is difficult to apply the standard Bayesian framework to represent agents who are in doubt about tautologies, but there is no reason to think that the correct decision theory is a tautology.

⁷Some reasons for worrying about whether information always has value do not arise in this context. We will not discuss cases where one might fail to rationally respond to new information (for discussion see Briggs 2009). The new information we are discussing is "transparent:" there are no relevant possibilities where you learn something but fail to know what it is that you have learned (see Das, forthcoming). We also set aside infinite cases (see Arntzenius and McCarthy 1997; Arntzenius, Elga, and Hawthorne 2004; Russell and Isaacs, forthcoming).

	EU, Heads	EU, Tails	REU, Heads	REU, Tails
Pass	32	32	32	32
Oracle	96	0	32	32

Which of these three options should Rita take? The answer depends on which decision theory is correct. By hypothesis, the correct decision theory is one of EU or REU, though Rita doesn't know which. So we can take them one by one, and in each case, the calculation is straightforward: see table 3.⁸

Table 3: The value that each decision theory assigns to each option

	EU	REU(p^2)
Gamble	48	24
Pass	32	32
Oracle	40	22

Both EU and REU agree on whether Rita should consult the oracle: *no*. The two theories give different verdicts on which option is best. EU ranks the three options Gamble > Oracle > Pass, and REU ranks them Pass > Gamble > Oracle. But both agree that Oracle is not the best option—in fact, both agree that Oracle is worse than Gamble. The risk-averse REU maximizer would rather not gamble, but by their lights consulting the oracle (and so *possibly* taking the gamble) is an even *worse* gamble than the original gamble.

If either of these two decision theories is correct, Rita should not consult the oracle. The basic reason for this is a structural feature that both of these theories have in common. Consulting the oracle amounts to choosing a *mixed strategy*: it has a chance of turning out the same way as Gamble, and a chance of turning out the same way as Pass, and which way it goes is independent of the outcomes of these acts. But EU and risk-averse REU

⁸Let's spell out one of these calculations:

$$\begin{aligned}
 \text{REU}(\text{Oracle}) &= 0 + r(3/4) \cdot (32 - 0) + r(1/4) \cdot (96 - 32) \\
 &= (3/4)^2 \cdot 32 + (1/4)^2 \cdot 64 = 18 + 4 \\
 &= 22
 \end{aligned}$$

both agree that a mixed strategy can be no better than the best of the pure strategies that comprise it.

The astute reader might wonder whether in the case of EU this result contradicts Good's Theorem, which says, in a slogan, that information is always valuable for EU-maximizers (Good 1966; see also Blackwell 1951). But the theorem only says that information is valuable for those who are *certain* to be EU-maximizers. This does not apply to Rita: she *might*, for all she knows, not turn out to be an EU-maximizer. For all she knows, if she asks the oracle then she will find out that EU is false, and instead be guided by REU.

This technical detail raises an important conceptual point. You might think that if Rita is sure to choose rationally (after consulting the oracle), and furthermore EU is the correct decision theory, then Rita is sure to maximize expected utility. But this doesn't follow: even if EU actually *is* the correct decision theory, it does not follow that it is *certain* to be the correct decision theory. Since Rita leaves open the possibility that some other decision theory is correct, she leaves open the possibility that she should not follow EU. In the "counternormative" situation where REU is correct instead of EU, Rita should Pass. If Rita is sure to do what she knows she should do, then she might not end up following EU.

At this point you might think: so much the worse for the idea that learning normative truths is always instrumentally valuable. That's one lesson you could draw, and it may turn out to be the right one in the end. But we shouldn't declare this idea defeated just yet. An alternative lesson to draw is that (contrary to the hypothesis of the example) neither EU nor REU(p^2) is the correct decision theory for Rita. In order to respect the idea that gaining normative information is worthwhile, we need a *new* decision theory: a theory of decisions under normative uncertainty.

To be clear, this decision theory would be a *rival* to standard EU and REU, not merely a neutral arbiter between them. Both EU and REU say Rita should not ask the oracle. If Rita *should* ask the oracle, then both of these theories are wrong. The correct decision theory is one we have not considered yet.

This sounds a bit different from how many people writing about normative uncertainty approach the issue (Sepielli 2009; 2018, 109–10; MacAskill 2016; MacAskill and Ord, forthcoming; Bykvist 2017). For instance, MacAskill (2016, 426) writes:

Proponents of metanormativism about moral uncertainty ...

think that there are different senses of 'ought': a first-order moral sense of 'ought,' which is not sensitive to a decision maker's moral uncertainty, and a different (more subjective or less idealized) sense of 'ought' that takes moral uncertainty into account.

The proposal is that "metanormative 'ought' " claims, which are sensitive to normative uncertainty, do not compete with "first-order 'ought' " claims that are not. There are various ways of explicating what the "metanormative 'ought' " might be (see Sepielli 2013; MacAskill, Bykvist, and Ord 2020, 20–21).

Here is how I prefer to frame things. What one ought to do depends on a body of information. When we talk about what you ought to do, normally we mean with respect to *your* current information, but not always; different bodies of information can be made salient in context.⁹ One can perfectly well raise to salience bodies of information that include all the normative facts. I will sometimes do this in what follows (though I will be explicit rather than relying on context). What one ought to do in light of all of the normative facts is "not sensitive to a decision maker's [normative] uncertainty," and we might say that this is what one ought to do "in a first-order sense." This contrasts with what one ought to do in light of just the information one actually has, which ordinarily does not include all the normative facts. I will be exploring the relationship between what one ought to do under normative uncertainty and what one ought to do given further normative information.

The theories EU and REU that I have been discussing are not "first-order" theories in this particular sense. They are theories of how what you should do depends on your information; these theories don't give a *special* place to normative uncertainty, but that isn't to say they don't take it into account (compare Harman 2015, 70). Indeed, these theories give normative uncertainty a non-trivial role. For example, suppose we offer Rita a simpler bet (compare Podgorski 2020, 45). We'll ask the oracle which decision theory is correct. If the oracle says that REU is correct, Rita gets 10 utils. If the oracle says EU is correct, Rita loses 2 utils. Should she take this bet? EU says yes: its expected utility is +4 utils. If EU was a "first-order" theory of what to do *given all of the normative facts*, it would not deliver this verdict. Conditional

⁹For clear discussion see Broome (1991, 126–29). Compare Kratzer ([1981] 2008); for critical discussion see Kolodny and MacFarlane (2010, 122–23); see also Pittard and Worsnip (2016)

on EU being correct, EU says Rita ought *not* take this bet: the conditional expected utility of the bet, given that EU is correct, is -2 utils. So what EU says Rita should do is sensitive to Rita’s uncertainty about whether EU is correct. Standard EU does not give any special *deference* to REU’s evaluations of acts, but that isn’t to say it is not sensitive to uncertainty about whether REU is correct. The question before us is *how* normative uncertainty is relevant to what we should do. EU and REU offer answers to this question—but if Rita should ask the oracle, then their answers are wrong.

I will approach this broad question by asking a narrower question: what kind of theory could deliver the verdict that learning the correct normative theory is instrumentally valuable?

2 Aggregating normative theories

Here is a natural idea (Nozick 1993; Lockhart 2000; Ross 2006; Sepielli 2009; MacAskill 2016; MacAskill and Ord, forthcoming). When we don’t know how well an act will actually turn out, it’s a good idea to aggregate the probabilities and utilities of different outcomes into a single summary value: the “choiceworthiness” of the act. EU and REU present different theories of choiceworthiness. Then, if we are uncertain about an act’s *actual choiceworthiness*, maybe it’s a good idea to consider its *expected choiceworthiness*. In Rita’s case, where two decision theories have equal probabilities, the obvious approach is to rank acts according to the average of their standard expected utility and their risk-weighted expected utility. In general, for any act A , define the *Expected Choiceworthiness*

$$EC(A) = \sum_i P(V_i \text{ is correct}) \cdot V_i(A)$$

where V_1, \dots, V_n are the value functions corresponding to each epistemically possible normative theory. (Here I make the simplifying assumptions that there are only finitely many possible theories, and that all of them assign acts values on the same quantitative scale. These assumptions will soon be relaxed.)

The theory that says to maximize Expected Choiceworthiness doesn’t advise consulting the oracle, either.¹⁰ This is easy to calculate (table 4). In fact,

¹⁰MacAskill (2014a; ch. 7; see also MacAskill, Bykvist, and Ord 2020, ch. 9) argues that moral information is instrumentally valuable in certain cases, taking Expected Choicewor-

the average of the EU of the Oracle act and its REU is *worse* than the average value for either Gamble or Pass.

Table 4: Values that Expected Choiceworthiness assigns to each option

	EU	REU(p^2)	EC
Gamble	48	24	36
Pass	32	32	32
Oracle	40	22	31

This example makes a broader structural point.¹¹ No decision theory that always recommends consulting normative oracles can be, in a natural sense, simply the result of aggregating the unconditional evaluations of epistemically possible decision theories like EU and REU in a way that satisfies **Naïve Valuewise Dominance**: if every epistemically live theory ranks A higher than B , then this decision theory also ranks A higher than B .¹² Rita’s case shows us that no way of aggregating theories that respects Naïve Valuewise Dominance will tell Rita to consult the normative oracle. Both of the live theories EU and REU rank Gamble over Oracle. So any aggregation rule that respects Naïve Valuewise Dominance, applied to these two theories, will *also* rank Gamble over Oracle.¹³

thiness maximization for granted. Rita’s case shows that this kind of argument does not generalize.

¹¹Podgorski (2020, 57) draws the same conclusion from a different, related argument.

¹²Besides EC, this applies to “My Favorite Theory,” (Gracely 1996; Gustafsson and Torpman 2014) “My Favorite Option” (see MacAskill and Ord, forthcoming), and the voting rules proposed by MacAskill (2016) and Tarsney (2019). (Note that this ordinal strict dominance principle is a bit different from the one discussed by MacAskill and Ord (forthcoming).)

¹³ There is another difficulty for EC. Before, we supposed that Rita knows the correct decision theory is either EU or REU—which implies that one of them is correct. But if Rita should consult the oracle, then neither is correct. Suppose we modify the case so that Rita assigns some positive credence to EC, in addition to EU and REU. Now the definition of EC is circular: EC itself appears as one of the terms on the right-hand side.

$$EC(A) = p_1 \cdot EU(A) + p_2 \cdot REU(A) + p_3 \cdot EC(A)$$

where p_1, p_2, p_3 are the probabilities of the three theories. But we can still make sense of this formula as a *constraint* on EC, which provides an *implicit* definition. As it happens, this gives exactly the same result as we get without accounting for the probability of EC:

$$EC(A) = \frac{p_1}{p_1 + p_2} \cdot EU(A) + \frac{p_2}{p_1 + p_2} \cdot REU(A)$$

That isn't to say the aggregation approach to normative uncertainty is totally off base. We just have to think carefully about what it is we should aggregate.

Standard EU doesn't think very highly of the Oracle act. But that's because it doesn't think very highly of how the Oracle act turns out in cases where EU is *not the correct decision theory*. We might say to EU: who asked you? Who cares what EU says about cases where EU is false? It would be irrational to be guided by EU in such cases!

Maybe it makes more sense to try to aggregate decision theories properly construed as *first-order* theories. Theories like EU and REU offer *conditional* evaluations of acts given various bodies of information. The "first-orderization" of a theory V like this consists of the evaluations that V offers *conditional on V being the correct decision theory*. The idea is that instead of trying to aggregate the "global" evaluations different possible decision theories offer—in effect, deferring to each theory's own account of decisions under normative uncertainty—we should instead aggregate just their "first-order" evaluations, in this sense. We care about what EU says about how to aggregate the values of cases where EU is correct; we don't have to listen to what it says about how to aggregate the values of other possible cases.¹⁴

Here is an example of a theory like this.¹⁵ Instead of taking the expected *unconditional* value delivered by each decision theory, instead we can consider the expected *conditional* value. For any act A , let

$$\text{ECC}(A) = \sum_i P(V_i \text{ is correct}) \cdot V_i(A \mid V_i \text{ is correct})$$

where V_1, \dots, V_n are the value functions corresponding to each epistemically possible decision theory. Call this **Expected Conditional Choicewor-**

(as long as $p_3 \neq 1$).

A similar issue arises for the Naïve Valuewise Dominance principle. If the correct normative theory is itself one of the live possibilities, then this principle is trivial: if every live theory ranks A over B , and the correct theory is one of those, then of course it ranks A over B . The principle that we must give up is that if every *other* live theory besides the correct one ranks A over B , then so does the correct theory. EC satisfies this principle as well.

¹⁴This way of thinking about "first-order" decision theories is partly inspired by the discussion of first-order and higher-order uncertainty by Dorst (2020).

¹⁵Podgorski (2020, 58) proposes essentially this theory; he calls his version "Maximize Inter-Theoretic Conditional Expectation (MITCE)." The relationship between EC and ECC is analogous to the contrast Elga (2013) draws between Rational Reflection and New Rational Reflection.

thiness, or ECC.¹⁶ This is obviously very similar in spirit to the Expected Choiceworthiness formula. In cases where the outcomes of the acts in question don't depend at all on which normative theory is correct, the two theories coincide. But the modified theory handles acts like Oracle very differently.

This theory *does* respect the value of normative information. Let's look at Rita's case again. The values that ECC assigns to Gamble and Pass are the same as EC (because the outcomes of Gamble and Pass are independent of which decision theory is correct). But the Oracle act goes differently, since its outcome depends on which decision theory is correct. If EU is correct, then the Oracle act goes the same way as Gamble. So

$$EU(\text{Oracle} \mid \text{EU is correct}) = EU(\text{Gamble} \mid \text{EU is correct}) = 48$$

If REU is correct, then the Oracle act goes the same way as Pass, instead. So

$$REU(\text{Oracle} \mid \text{REU is correct}) = REU(\text{Pass} \mid \text{REU is correct}) = 32$$

So $ECC(\text{Oracle}) = 40$ (the average of 48 and 32). This is better than either Gamble or Pass, as we hoped (table 5). The basic reason for this is that the Oracle act gets the best of both worlds: its score is the average of the EU score for the best EU-scoring act and the REU score for the best REU-scoring act. This has to be at least as good as the average of any single act's EU score and its REU score.

Table 5: Values that Expected Conditional Choiceworthiness assigns to each option

	EU	REU(p^2)	EC	ECC
Gamble	48	24	36	36
Pass	32	32	32	32
Oracle	40	22	31	40

¹⁶The issue raised in footnote 13 arises again: if ECC is itself one of the epistemically possible theories, then this would be a circular definition. But it still provides a perfectly sensible *constraint* on a decision theory—we should read it this way, rather than as a complete theory of its own. This tells us the unconditional values of an act in terms of the *conditional*, “first-order” values each theory assigns the act—including the correct first-order theory, whatever it may be.

This holds in general: ECC always permits consulting normative oracles when they are available. (ECC *requires* consulting normative oracles whenever this might make a difference to one’s eventual choice.) There are also other rules of this same general form that use different aggregation methods besides the simple expectation: for example, we could have used a “second-order” risk-weighted expectation instead of the standard expectation. Rules like this also recommend consulting normative oracles. What’s more, a kind of converse also holds: if normative information is valuable, decision theory under normative uncertainty must take this general form. In the rest of this section I will spell out this idea precisely and explain some of its consequences.

First a bit of technical set-up. Consider a finite set \mathcal{S} of **states**, and a set \mathcal{O} of **outcomes**.¹⁷ A **proposition** is a set of states. **Acts** are functions from states to outcomes. For an act A and a proposition E , let the **partial act** $A|E$ be the restriction of A to E .

Let a **normative theory** be a function that takes each non-null proposition E to an ordering of acts \succsim_E . Intuitively, $A \succsim_E B$ means that act A is at least as good a choice as act B given the information E . (The unconditional ordering $A \succsim B$ is $A \succsim_{\mathcal{S}} B$.) We will assume that each normative theory’s ordering \succsim_E only depends on what happens when E obtains: if $A|E = B|E$ then $A \sim_E B$.¹⁸ We will also assume these orderings satisfy some standard technical conditions.¹⁹

For each state $s \in \mathcal{S}$, let \succsim_E^s be the normative theory which is **correct at** s . For each $s \in \mathcal{S}$, let $V(s)$ intuitively represent the proposition that the s -theory is correct: this is the set of states s' such that \succsim_E^s and $\succsim_E^{s'}$ agree

¹⁷Intuitively, we will think of each state as having positive probability. If some states in \mathcal{S} don’t represent genuinely “live” possibilities, then the principles that follow require qualification.

¹⁸This principle is sometimes called *consequentialism*, but it is not the only such principle; note that this principle is compatible with risk-aversion (for discussion see Hammond 1988; Machina 1989; Buchak 2013, ch. 6).

¹⁹ In particular, the conditions for Debreu’s ordinal utility representation theorem (1954, 1959): the set of outcomes has the structure of a connected separable topological space, and each of the orderings is transitive, total, and continuous. These technical assumptions simplify our discussion, by allowing us to treat each theory as assigning values that are at least ordinally representable by *real numbers*. But note that we are effectively supposing that *no epistemically live theory* allows intransitivity, incompleteness, or infinite utilities. This is a strong assumption. I do not think that the core lessons of this paper depend on it, but dropping it will lead us to consider non-standard value representations—such as vector utility and multi-utility representations. These complications are best left for later.

(for every E). Let $s^* \in \mathcal{S}$ be the **actual state**. We'll use the shorthand \succsim^* for the ordering \succsim^{s^*} .

When Rita consults the decision theory oracle, she learns which normative theory is correct. She will then choose an act recommended by that theory. But in general this may not be the act that the theory recommended *before* she consulted the oracle. She has new information now, and this can make a difference to what she should choose. So she will choose an act which is recommended by the theory that turns out to be correct, *given this new information* that this theory is correct.

Let \mathcal{A} be any finite set of acts. Call an act A **\mathcal{A} -oracular** iff for each state s , there is some act $B \in \mathcal{A}$ such that $A|V(s) = B|V(s)$ and

$$B \succsim_{V(s)}^s B' \quad \text{for each act } B' \in \mathcal{A}$$

Intuitively, this is an act that consists in choosing whichever act is recommended by the correct normative theory, given the information that this theory is correct. Our principle is that an act like this is at least as good as any of the originally available acts. A stronger version is that consulting the oracle is strictly *better* than committing to an act in advance, unless there is already an available act which we know in advance will be the best option no matter which theory turns out to be true.

Value of Normative Information. Let \mathcal{A} be any set of acts and let A be any act.

(Weak) For each act $B \in \mathcal{A}$, if A is \mathcal{A} -oracular, then $A \succsim^* B$

(Strong) If furthermore B is *not* \mathcal{A} -oracular, then $A \succ^* B$.

Valuewise Dominance. Let A and B be any acts.

(Weak) If $A \succsim_{V(s)}^s B$ for each state s , then $A \succsim^* B$.

(Strong) If furthermore there is some state s such that $A \succ_{V(s)}^s B$, then $A \succ^* B$.

A function u from acts to real numbers (**ordinally**) **represents** an ordering \succsim of acts iff

$$A \succsim B \quad \text{iff} \quad u(A) \geq u(B) \quad \text{for any acts } A \text{ and } B$$

In the case of EU and REU, we specified the orderings by way of real-valued functions that represent those orderings.

Aggregation. Let $\succsim_1, \succsim_2, \dots$ list all of the distinct orders $\succsim_{V(s)}^s$ for each state s . There are functions u_1, u_2, \dots such that each u_i represents \succsim_i , and there is a function f from tuples of real numbers to real numbers, such that the function that takes each act A to $f(u_1(A), u_2(A), \dots)$ ordinally represents the unconditional ordering \succsim^* . Furthermore:

(Weak) If $x_i \geq y$, then

$$f(x_1, \dots, x_i, \dots) \geq f(x_1, \dots, y, \dots)$$

(Strong) If $x_i > y$, then

$$f(x_1, \dots, x_i, \dots) > f(x_1, \dots, y, \dots)$$

Proposition 1.

- (a) *Weak Value of Normative Information, Weak Valuewise Dominance, and Weak Aggregation are equivalent.*
- (b) *Strong Value of Normative Information, Strong Valuewise Dominance, and Strong Aggregation are equivalent.*²⁰

Aggregation tells us that normative theories can be split up into two levels: a “first-order” theory and a “second-order” theory, which are autonomous from one another. The *first-order* part of a normative theory tells us how

²⁰*Proof Sketch.*

Value of Normative Information implies Valuewise Dominance. Let A and B be acts such that $A \succsim_{V(s)}^s B$ for each state s . Then A is $\{A, B\}$ -oracular. So Weak Value of Normative Information implies $A \succsim^* B$. If furthermore $A \succ_{V(s)}^s B$ for some state s , then B is not $\{A, B\}$ -oracular. So Strong Value of Normative Information implies $A \succ^* B$.

Valuewise Dominance implies Value of Normative Information. Let A be \mathcal{A} -oracular, and let $B \in \mathcal{A}$. So $A \succsim_{V(s)}^s B$ for each state s , and Weak Valuewise Dominance implies $A \succsim^* B$. If furthermore B is not \mathcal{A} -oracular, then $A \succ_{V(s)}^s B$ for some state s , and so $A \succ^* B$.

Valuewise Dominance is equivalent to Aggregation. Debreu (1954, Theorem I) tells us that (given the technical assumptions in footnote 19) each of the orders \succsim_i is representable by a utility function u_i , and also that \succsim^* is representable by a utility function U . Weak Valuewise Dominance then tells us that, for any acts A and B , if $u_i(A) \geq u_i(B)$ for each i , then $U(A) \geq U(B)$. Applying this in both directions, if $u_i(A) = u_i(B)$ for each i , then $U(A) = U(B)$. So there is a function f such that $U(A) = f(u_1(A), u_2(A), \dots)$. It is clear that Weak/Strong Valuewise Dominance implies that f is weakly/strongly increasing in each argument. The converse is clear. (Note that the “Strong” case follows from Gorman (1968, Theorem 2).)

it orders acts conditional on its own correctness. The *second-order* part of the theory is its unconditional ordering; for a theory that satisfies Value of Normative Information this can be represented with an aggregation function f that combines the verdicts of different possible first-order theories. The actually correct first-order theory supplies one argument to this aggregation function, and otherwise plays no special role in determining what you should do *unconditionally*.

These two pieces are separate: you can pick any first-order theory and combine it with any aggregation function, so far as anything we've said so far goes. For example, there can be a decision theory that agrees with the first-order judgments of EU (that is, the conditional ordering $\succsim_{V(s)}^s$ is represented by the conditional expected utility function $EU(- | V(s))$) and whose aggregation function is given by $REU(p^2)$, giving extra weight to theories that evaluate an act as worse. There is another decision theory that works the other way around. Both of these combinations satisfy Value of Normative Information. (For general discussion of metanormative risk-aversion, see Dietrich and Jabarian, manuscript.)

While Aggregation is permissive, it imposes some real constraints. First, we have already noted that it is incompatible with simple Expected Choiceworthiness—it requires that we aggregate *conditional* evaluations. It also rules out both of the following principles:

First-Order Actualism. $A \succsim^* B$ iff $A \succsim_{V(s^*)}^* B$, for all acts A, B .

My Favorite First-Order Theory. $A \succsim^* B$ iff $A \succsim_{V(s)}^s B$, for all acts A, B , where s is a state such that $V(s)$ has the highest probability.

Both of these principles violate Strong Valuewise Dominance. If the actually correct first-order theory (or the most probable first-order theory) is indifferent between A and B , then A and B will be indifferent unconditionally even if all the other live theories break the tie in the same way. Strong Aggregation requires that we give every live theory at least a *little* say—at least as a tie-breaker.²¹

proposition 1 has nothing special to do with risk-aversion. The live normative theories in question could be any ways of ranking acts, and they could

²¹ Compare MacAskill and Ord (forthcoming, 7–8). Note that Aggregation is compatible with *lexical* aggregation, where there is some “importance” ordering on theories, and each theory’s verdicts *only* make a difference as tie-breaking consideration when all of the more important theories have ties. (Gustafsson and Torpman 2014 defend a theory with this structure.)

differ in many different important respects. We have considered theories that disagree about how to weigh outcomes, but this framework also applies to theories that disagree about how valuable different particular outcomes are. How bad is animal suffering? How good are large populations? It seems plausible that it would be instrumentally morally valuable to know the answers to such questions. If this is so, then we should not just act so as to promote whatever is *actually* good (because Strong Value of Normative Information conflicts with First-Order Actualism); rather we should aggregate the various live possible first-order theories of the good in a way that gives each of them at least a bit of say.

3 Additive Aggregation

So far we have considered consequences of the principle that you should learn the *completely specific* normative truth, given the opportunity. A stronger principle suggests itself: you should also take opportunities to learn less complete normative truths. This principle requires that the aggregation function takes a more specific form: in fact, it pushes us pretty close to Expected Conditional Choiceworthiness.

A **normative proposition** N is, intuitively, a proposition entirely about the question of which normative theory is correct: for any state s , if $V(s)$ is consistent with N then $V(s)$ entails N . We'll continue to consider cases that are idealized enough that you know what you would do with new information if you got it. But now we *won't* assume that you know you will do what you *should* do: if N provides only incomplete normative information, then you may not know what you should do, and you might do something else instead. Let $A \succsim^N B$ mean that, given the normative information N (and nothing else), you would choose act A over B . This ordering should satisfy the same technical conditions as before (footnote 19). In particular, we assume it is "consequentialist" in the sense that for any acts A and B , if $A|N = B|N$ then $A \sim^N B$. We assume that in the case where $N = V(s)$ for some state s , this agrees with what we said before: $A \succsim^N B$ iff $A \succsim_{V(s)}^s B$. Then the general principle that it is valuable to learn whether a normative proposition N is true can be written as another dominance-style principle.

Value of General Normative Information. For any normative proposition N ,

(Weak) If $A \succsim^N B$ and $A \succsim^{-N} B$, then $A \succsim^* B$.

(Strong) Furthermore, if $A \succ^N B$ and $A \succsim^{-N} B$, then $A \succ^* B$.

This is a strong constraint on decision theory under normative uncertainty. In fact, given our other background assumptions it requires that this theory is something in the same family as Expected Conditional Choiceworthiness.

Additive Aggregation. There are utility functions u_1, u_2, \dots from acts to real numbers such that each distinct ordering $\succsim_{V(s)}^s$ is represented by some function u_i , and the sum

$$U(A) = u_1(A) + u_2(A) + \dots$$

represents \succsim^* .

Proposition 2. *Strong Value of General Normative Information and Additive Aggregation are equivalent.*²²

Additive Aggregation includes Expected Conditional Choiceworthiness as a special case, but it is less constrained. First, while it permits weighing theories by their probabilities (by appropriately scaling the utility function u_i) this is not required.²³ More importantly, the utility functions u_i don't have to match any "internal" cardinal choiceworthiness functions the theories themselves might deploy, like standard expected utilities or REU values. They are only required to represent the *ordinal* rankings of acts delivered by those theories. The live theories don't even have to assign *any* cardinal utilities to acts "internally"—the framework only relies on ordinal rankings. And unlike ECC, Additive Aggregation doesn't depend on any pre-established "inter-theoretic comparisons of value."²⁴

²²Value of General Normative Information implies that each normative proposition is *separable* (see section 4). Then (Gorman 1968, Theorem 2) implies that the unconditional order of acts is represented by a sum of subutilities which are defined on the "atomic" normative propositions, namely the propositions $V(s)$; these subutility functions represent the orders $\succsim_{V(s)}^s$.

²³The official framework laid out in section 2 is one of (unquantified) "uncertainty" rather than (quantified) "risk"—we have not actually assumed that any probabilities are given at all.

²⁴Such comparisons are one of the standard difficulties "metanormative" decision theories face (Hedden 2016; MacAskill 2014b; Ross 2006; Sepielli 2009; Carr, forthcoming). Note that while our framework didn't take any such comparisons for granted, Additive Aggregation *imposes* structure that allows some interesting intertheoretic cardinal comparison (compare Riedener 2019, 2020; see also Dietrich and Jabarian, manuscript). The functions u_1, u_2, \dots are unique up to positive affine transformations with the same scaling constant

While Additive Aggregation does not push us to a unique aggregation rule, it is constrained enough to rule out many theories. For example, the risk-weighted expectation is not additively separable in this way.²⁵

4 Normative information or empirical information?

You might think that learning the correct normative theory is instrumentally valuable because learning new information is instrumentally valuable across the board. This natural thought is untenable. In fact, always valuing normative information is inconsistent with always valuing empirical information.²⁶

To be more precise, this inconsistency holds as long as it is epistemically possible that risk-aversion is rational. The basic reason is that risk-averse decision theories like $REU(p^2)$ sometimes recommend avoiding empirical information (see Wakker 1988; Buchak 2010; 2013, 171ff.; Briggs 2016; Campbell-Moore and Salow, forthcoming; Salow, forthcoming). New information can be misleading: it can lead you to make a decision that unluckily turns out worse than the one you would have made otherwise. In standard risk-neutral decision theory, the risk of being thus misled by new information is always balanced off by the ways in which the more informed decision can turn out better. But risk-averse agents may not regard this trade-off as acceptable.

If you think risk-aversion *might* be rational, and you always respect the value of normative information, then there are possible circumstances in which you prefer to avoid empirical information. But it turns out that this merely *possible* information-avoidance also infects the *actual* value of empirical information. If information *might* not always be valuable, then it *isn't* always valuable.

This form of argument is familiar from the moral uncertainty literature. Lockhart (2000) argues that if abortion *might* be impermissible, then it should not be performed. Guerrero (2007) argues that if animals *might* have moral status, then we should not kill them. MacAskill, Bykvist,

(see Broome 1991, 74–75). This standardizes a way of trading off gains in utility given one first-order normative theory against losses in utility given another.

²⁵For any risk function other than $r(p) = p$, which reduces to the standard expectation. Note that Additive Aggregation also rules out lexical aggregation (footnote 21).

²⁶The argument in this section only relies on the Value of (Specific) Normative Information principle from section 2, not the Value of General Normative Information from section 3.

and Ord (2020, ch. 8) object to these arguments; but they go on to offer several structurally parallel arguments of their own. If it *might* be that you should give special weight to the well-being of friends over that of strangers, then it really *is* appropriate to give them some (lesser) special weight.²⁷ If it *might* be that you should prioritize the worst off, then it *is* appropriate to prioritize the worst off to some extent. If there *might* be objective goods besides people’s welfare, then it actually *is* appropriate to promote those other considerations. And so on.²⁸ The argument in this section has a similar form: it says that if it *might* be rational to avoid information, then it *is* rational to avoid information—either empirical or normative. One distinctive feature of this argument is that it relies on much weaker assumptions about normative uncertainty than the other arguments just mentioned—we need not take for granted anything as strong as maximizing expected choiceworthiness.

We can show the conflict with another example. Rico is undecided between two decision theories, EU_1 and REU_1 . The first-order evaluations of EU_1 agree with standard EU theory: that is, EU_1 ranks $A \succ B$ iff $EU(A \mid EU_1 \text{ is correct}) \geq EU(B \mid EU_1 \text{ is correct})$. The first-order evaluations of REU_1 agree with $REU(p^2)$, in the same sense. Rico faces a more complicated gamble. Its pay-off depends on the outcome of two fair coin flips *and* which decision theory is correct, as in table 6. (The decision theory oracle has agreed to help settle the bet.)

Table 6: Utilities of the outcomes for a more complicated gamble

	EU_1	REU_1
Heads, Heads	23	96
Heads, Tails	23	0
Tails, Heads	23	32
Tails, Tails	23	32

Rico faces the following choice: he can commit to *Gamble*, or he can commit to *Pass* and keep utility 23 no matter what. But he also has a third option:

²⁷“Appropriate” is their preferred term for the actions that are recommended by a theory of moral uncertainty.

²⁸These arguments are also reminiscent of the modal ontological argument: if a necessary being *could* exist, it *does* exist.

Learn how the first coin landed, and then choose whether to Gamble or Pass.

If empirical information has instrumental value, then Rico should choose to Learn. But we can show that the value of *normative* information implies that Rico should commit to Pass instead. For any act A (or partial act), let $EU_1(A)$ denote the “first-order” value $EU(A \mid EU_1 \text{ is correct})$, and similarly let $REU_1(A) = REU(A \mid REU_1 \text{ is correct})$.

First, it is easy to calculate:

$$\begin{aligned}EU_1(\text{Gamble}) &= 23 \\REU_1(\text{Gamble}) &= 22 < 23\end{aligned}$$

So by Strong Valuewise Dominance, $\text{Pass} >^* \text{Gamble}$.

Second, we can show that after Rico finds out how the coin lands he should choose to Gamble, either way. Let H be the proposition that the first coin lands Heads. Then again we can calculate:

$$\begin{aligned}EU_1(\text{Gamble} \mid H) &= 23 \\REU_1(\text{Gamble} \mid H) &= 24 > 23 \\EU_1(\text{Gamble} \mid \neg H) &= 23 \\REU_1(\text{Gamble} \mid \neg H) &= 32 > 23\end{aligned}$$

So by Strong Valuewise Dominance, first for values conditional on H , and then for values conditional on $\neg H$, we have:

$$\begin{aligned}\text{Gamble} &>_H^* \text{Pass} \\ \text{Gamble} &>_{\neg H}^* \text{Pass}\end{aligned}$$

Either way, Rico should choose to Gamble.

Rico knows this; so Rico knows that Learn will have the same outcome as Gamble, no matter what. So Learn is strictly *worse* than committing to Pass without getting the information about the coin flip. Rico should Pass rather than Learn, avoiding empirical information.

Moral. Valuing *normative* information conflicts with valuing *empirical* information.

Here is a more general argument for this moral. (The following draws on ideas from Broome 1991, ch. 4; Debreu 1959; Gorman 1968.) In general, value of information principles imply corresponding *separability* principles.

A proposition E is **separable** (for an ordering of acts \succsim) iff there exists an ordering \succsim_E of *partial* acts defined on E such that, for any acts A and B whose restrictions to $\neg E$ are the same ($A|\neg E = B|\neg E$), we have $A \succsim B$ iff $A|E \succsim_E B|E$.

If empirical information has value, this tells us that each empirical proposition E is separable. Given Strong Value of Normative Information, this implies that E is also separable for each *first-order* normative theory. In general, though, empirical propositions are *not* separable for risk-averse theories like REU. So this requires that no such theories are live possibilities.

Proposition 3. *Given Strong Value of Normative Information, if E is separable for \succsim^* then for each state s , E is separable for $\succsim_{V(s)}^s$.*²⁹

We might also suppose that the empirical propositions *cross-cut* the normative propositions: each state is determined by an empirical proposition together with a normative proposition. In that case, the two value of information principles together imply that *every* proposition is separable. This implies that the ordering is in fact additively separable over states: if the states are s_1, s_2, \dots , there are functions u_1, u_2, \dots from outcomes to real numbers, such that \succsim^* is represented by the function that takes each act A to

$$u_1(A(s_1)) + u_2(A(s_2)) + \dots$$

(Again see Broome 1991, ch. 4.) This *statewise* additivity is a stronger constraint than *normative-theory-wise* Additive Aggregation principle discussed in section 4. It amounts to requiring that the only live normative theories are standard expected utility at the first order, with only the choice of probabilities and utilities left open (though the utility functions may be “state-dependent”), and that these first-order expected utilities are then themselves aggregated simply by adding them up.

If any live first-order theories are non-additive, we have a conflict between two very similar, and similarly compelling principles. Which should go?

Maybe it’s time to give up on the Value of Normative Information: if empirical information really does have value, but we are *uncertain* about whether risk-aversion is rational, and thus about whether empirical information has value, then we should not always consult normative oracles. Moreover, this implies that Aggregation is false. In general, we cannot consistently value information while also simultaneously deferring (even a little, even as a

²⁹This follows from results in Gorman (1968).

tie-breaker) to theories that do not value information. Thus a standard “dynamic consistency” argument for expected utility theory becomes an argument *against* aggregative theories of normative uncertainty.

Alternatively, we might run this argument the other way: maybe normative information takes precedence over empirical information. It may still be true that the correct *first-order* normative theory respects empirical information. Perhaps, when the actual normative facts are known, we should maximize EU. But if for all we know the correct normative theory might *not* respect the value of empirical information, then we ought to hedge.

Either one of these conclusions would teach an important abstract lesson: uncertainty about normative principles is structurally different from empirical uncertainty. (This echoes Weatherson 2019, 39.) The two kinds of uncertainty do not have the same significance for information-gathering.³⁰

One other possibility is that each principle undermines the support for the other. We might take this conflict to point us toward more permissive normative theories like REU that sometimes recommend avoiding information. Rationally avoiding information is weird.³¹ But perhaps it is unavoidable.

I don’t know which of these lessons is correct: I find the conflict between the value of normative information and the value of empirical information quite puzzling.

5 Metanormative uncertainty

There is a standard regress objection to the project of theorizing about normative uncertainty (Weatherson 2014, 155–57). Whatever the correct “metanormative” theory is, couldn’t it be doubted too? And doesn’t this call for a theory of decisions under “metanormative uncertainty”—and so on indefinitely?

In our framework, a normative theory V includes a “first-order” theory— V ’s prescriptions conditional on the correctness of V —and a “second-order”

³⁰I speak of “normative principles” here because the argument leaves open that empirical information and *some* kinds of normative information both have value. While statewise additivity rules out deferring to risk-averse decision theories, it is compatible with deferring to alternative theories of the *value of outcomes*. (Consider the “ex post expectationism” discussed by Dietrich and Jabarian, manuscript, 12; compare also MacAskill, Bykvist, and Ord 2020, 107–8; Tarsney, Manuscript.)

³¹Even Buchak (2013, 199) concedes that it is “somewhat unpalatable.”

part, which gives V 's unconditional prescriptions. Uncertainty about either one of these parts is uncertainty about whether V is correct. "Metanormative" uncertainty is thus a kind of normative uncertainty.

Aggregative theories (as in section 2) say that there is some aggregation function, and some utility functions representing each *first-order* normative theory, and you should do whatever maximizes the aggregate value. And if you don't know which function that is? Aggregative theories don't care. They give no place to any kind of "second-order aggregation," which would somehow aggregate different aggregation functions. What you should do (unconditionally) is determined by the *actually correct* aggregation function together with the live normative theories' *first-order* evaluations, conditional on their own correctness. Other kinds of normative uncertainty don't make any difference to what you should do.

What about the value of "metanormative" information? Simple reasoning by analogy might suggest that just as preferring to learn the answers to normative questions led to a theory that aggregates first-order normative theories, preferring to learn the answers to *metanormative* questions—about which aggregation function is correct—would lead to a theory of "second-order aggregation." But it doesn't work out that way. In our framework, "metanormative" questions are themselves normative questions. *Additive* theories (as in section 3) recommend learning arbitrary normative information when the opportunity arises, and so, in particular, such theories recommend learning "metanormative" information. This is despite the fact that additive theories don't include any "second-order aggregation" of different possible aggregation functions. In *this* respect, the Value of Normative Information pushes us toward the side of the "actualist" (Harman 2015; Weatherston 2014, 2019; Hedden 2016). The *correct* theory of what you should do is the correct theory of *what you should do*.

You can be genuinely uncertain about you should do. The value of normative information leads to the view that what you should do is determined by some way of aggregating the evaluations of competing epistemically possible first-order theories. But this was not motivated by the idea that there is some sense in which you *can* tell what you "really" should do in some more "action-guiding" sense. (Contrast Sepielli 2018, 113; MacAskill and Ord, forthcoming, 5.) The question of what to do when we face normative uncertainty is not hostage to the presumption that what we should do is always knowable (for critical discussion of this idea see Srinivasan 2015). Tragically often there is no way of knowing for sure what you should do. If

it is instrumentally valuable to learn normative truths, then this uncertainty matters for what you actually *should* do. But that does not provide any reason for optimism that uncertainty about we should do can be escaped.

References

- Allais, M. 1953. "Le Comportement de l'homme Rationnel Devant Le Risque: Critique Des Postulats Et Axiomes de l'ecole Americaine." *Econometrica* 21 (4): 503–46. <https://doi.org/10.2307/1907921>.
- Arntzenius, Frank, Adam Elga, and John Hawthorne. 2004. "Bayesianism, Infinite Decisions, and Binding." *Mind* 113 (450): 251–83.
- Arntzenius, Frank, and David McCarthy. 1997. "The Two Envelope Paradox and Infinite Expectations." *Analysis* 57 (1): 42–50.
- Blackwell, David. 1951. "Comparison of Experiments." In. The Regents of the University of California. <https://projecteuclid.org/euclid.bsm/1200500222>.
- Briggs, R. A. 2009. "Distorted Reflection." *Philosophical Review* 118 (1): 59–85.
- . 2016. "Costs of Abandoning the Sure-Thing Principle." *Canadian Journal of Philosophy* 45 (5): 827–40.
- Broome, John. 1991. *Weighing Goods: Equality, Uncertainty and Time*. Wiley-Blackwell.
- Buchak, Lara. 2010. "Instrumental Rationality, Epistemic Rationality, and Evidence-Gathering." *Philosophical Perspectives* 24 (1): 85–120. <https://doi.org/10.1111/j.1520-8583.2010.00186.x>.
- . 2013. *Risk and Rationality*. Oxford University Press.
- Bykvist, Krister. 2017. "Moral Uncertainty." *Philosophy Compass* 12 (3): e12408.
- Campbell-Moore, Catrin, and Bernhard Salow. Forthcoming. "Avoiding Risk and Avoiding Evidence." *Australasian Journal of Philosophy*, 1–21. <https://doi.org/10.1080/00048402.2019.1697305>.
- Carr, Jennifer Rose. Forthcoming. "Normative Uncertainty Without Theories." *Australasian Journal of Philosophy*, 1–16. <https://doi.org/10.1080/00048402.2019.1697710>.
- Das, Nilanjan. Forthcoming. "The Value of Biased Information." *British Journal for the Philosophy of Science* 0: axaa003.
- Debreu, Gerard. 1954. "Representation of a Preference Ordering by a Numerical Function." *Decision Processes* 3: 159–65.
- . 1959. "Topological Methods in Cardinal Utility Theory." Cowles Foundation for Research in Economics, Yale University.

- Dietrich, Franz, and Brian Jabarian. Manuscript. "Axiomatic of Normative Uncertainty."
- . Manuscript. "Expected Value Under Normative Uncertainty."
- . Manuscript. "The Risk Attitude Under Normative Uncertainty."
- Dorst, Kevin. 2020. "Evidence: A Guide for the Uncertain." *Philosophy and Phenomenological Research* 100 (3): 586–632. <https://doi.org/10.1111/phpr.12561>.
- Elga, Adam. 2013. "The Puzzle of the Unmarked Clock and the New Rational Reflection Principle." *Philosophical Studies* 164 (1): 127–39. <https://doi.org/10.1007/s11098-013-0091-0>.
- Good, I. J. 1966. "On the Principle of Total Evidence." *British Journal for the Philosophy of Science* 17 (4): 319–21.
- Gorman, W. M. 1968. "The Structure of Utility Functions." *Review of Economic Studies* 35 (4): 367–90. https://econpapers.repec.org/article/ouprestud/v_3a35_3ay_3a1968_3ai_3a4_3ap_3a367-390..htm.
- Gracely, Edward J. 1996. "On the Noncomparability of Judgments Made by Different Ethical Theories." *Metaphilosophy* 27 (3): 327–32. <https://doi.org/10.1111/j.1467-9973.1996.tb00212.x>.
- Guerrero, Alexander A. 2007. "Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution." *Philosophical Studies* 136 (1): 59–97.
- Gustafsson, Johan E., and Olle Torpman. 2014. "In Defence of My Favourite Theory." *Pacific Philosophical Quarterly* 95 (2): 159–74. <https://doi.org/10.1111/papq.12022>.
- Hammond, Peter J. 1988. "Consequentialist Foundations for Expected Utility." *Theory and Decision* 25 (1): 25–78. <https://doi.org/10.1007/bf00129168>.
- Harman, Elizabeth. 2015. "The Irrelevance of Moral Uncertainty." *Oxford Studies in Metaethics* 10.
- Hedden, Brian. 2016. "Does MITE Make Right? Decision-Making Under Normative Uncertainty." In *Oxford Studies in Metaethics Volume 11*, edited by Russ Schafer-Landau, 102–28.
- Hudson, James L. 1989. "Subjectivization in Ethics." *American Philosophical Quarterly* 26 (3): 221–29.
- Kolodny, Niko, and John MacFarlane. 2010. "Ifs and Oughts." *Journal of Philosophy* 107 (3): 115–43. <https://doi.org/10.5840/jphil2010107310>.
- Kratzer, Angelika. (1981) 2008. "The Notional Category of Modality." In *Formal Semantics: The Essential Readings*, edited by Paul H. Portner and Barbara H. Partee, 35. John Wiley & Sons, Incorporated.
- Lockhart, Ted. 2000. *Moral Uncertainty and Its Consequences*. Oxford University Press.

- MacAskill, William. 2014b. "Normative Uncertainty." PhD thesis, University of Oxford.
- . 2014a. "Normative Uncertainty." PhD thesis, Oxford University.
- . 2016. "Normative Uncertainty as a Voting Problem." *Mind* 125 (500): 967–1004.
- MacAskill, William, Krister Bykvist, and Toby Ord. 2020. *Moral Uncertainty*. MacAskill, William, and Toby Ord. Forthcoming. "Why Maximize Expected Choice-Worthiness?" *Noûs*.
- Machina, Mark J. 1989. "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty." *Journal of Economic Literature* 27 (4): 1622–68. <http://www.jstor.org/stable/2727025>.
- Nozick, Robert. 1993. *The Nature of Rationality*. Princeton University Press.
- Pittard, John, and Alex Worsnip. 2016. "Metanormative Contextualism and Normative Uncertainty." *Mind* 126 (501): 155–93. <https://doi.org/10.1093/mind/fzv208>.
- Podgorski, Abelard. 2020. "Normative Uncertainty and the Dependence Problem." *Mind* 129 (513): 43–70.
- Quiggin, John. 1982. "A Theory of Anticipated Utility." *Journal of Economic Behavior & Organization* 3 (4): 323–43. [https://doi.org/10.1016/0167-2681\(82\)90008-7](https://doi.org/10.1016/0167-2681(82)90008-7).
- . 1993. *Generalized Expected Utility Theory: The Rank-Dependent Model*. Springer Netherlands. <https://doi.org/10.1007/978-94-011-2182-8>.
- Riedener, Stefan. 2019. "Constructivism about Intertheoretic Comparisons." *Utilitas* 31 (3): 277–90. <https://doi.org/10.1017/s0953820819000165>.
- . 2020. "An Axiomatic Approach to Axiological Uncertainty." *Philosophical Studies* 177 (2): 483–504. <https://doi.org/10.1007/s11098-018-1191-7>.
- Ross, Jacob. 2006. "Rejecting Ethical Deflationism." *Ethics* 116 (4): 742–68.
- Russell, Jeffrey Sanford, and Yoav Isaacs. Forthcoming. "Infinite Prospects." *Philosophy and Phenomenological Research* 0.
- Salow, Bernhard. Forthcoming. "The Value of Evidence." In *The Routledge Handbook of the Philosophy of Evidence*, edited by Maria Lasonen-Aarnio and Clayton M. Littlejohn. Routledge.
- Sepielli, Andrew. 2009. "What to Do When You Don't Know What to Do." *Oxford Studies in Metaethics* 4: 5–28.
- . 2013. "What to Do When You Don't Know What to Do When You Don't Know What to Do..." *Noûs* 47 (1): 521–44.
- . 2018. "How Moral Uncertainty Can Be Both True and Interesting." *Oxford Studies in Normative Ethics* 7.
- Srinivasan, Amia. 2015. "Normativity Without Cartesian Privilege." *Philo-*

- sophical Issues* 25 (1): 273–99. <https://doi.org/10.1111/phis.12059>.
- Tarsney, Christian. 2019. “Normative Uncertainty and Social Choice.” *Mind* 128 (512): 1285–1308.
- . Manuscript. “Metanormative Regress: An Escape Plan.”
- Wakker, Peter. 1988. “Nonexpected Utility as Aversion of Information.” *Journal of Behavioral Decision Making* 1 (3): 169–75. <https://doi.org/10.1002/bdm.3960010305>.
- Weatherson, Brian. 2014. “Running Risks Morally.” *Philosophical Studies* 167 (1): 141–63. <https://doi.org/10.1007/s11098-013-0227-2>.
- . 2019. *Normative Externalism*. Oxford University Press.