

# Introduction

We are often uncertain about what we ought, morally, to do. Suppose that Alice has £20 to spend. With that money, she could eat out at a pleasant restaurant. Alternatively, she could pay for four long-lasting insecticide-treated bed nets that would protect eight children against malaria for two years.<sup>1</sup> Let's suppose that Alice knows all the morally relevant empirical facts about what that £20 could do. Even so, it might be that she still doesn't know whether she's obligated to donate that money or whether it's permissible for her to pay for the meal out, because she just doesn't know how strong her moral obligations to distant strangers are. If so, then even though Alice knows all the relevant empirical facts, she doesn't know what she ought to do.

Or suppose that the members of a government are making a decision about whether to tax carbon emissions. Let's assume that they know all the relevant facts about what would happen as a result of the tax: it would make presently existing people worse off, since they would consume less oil and coal, and would therefore be less economically productive; but it would slow down climate change, thereby on balance increasing the welfare of people living in the future. But the members of the government don't know how to weigh the interests of future people against the interests of presently existing people. So, again, the members of this government don't ultimately know what they ought to do.

These are instances of *moral uncertainty*: uncertainty that stems not from uncertainty about descriptive matters, but about moral or evaluative matters. Moral uncertainty is commonplace: given the difficulty of ethics and the widespread disagreement about ethical issues, moral uncertainty is not the exception, but the norm.

Moral uncertainty matters. If we don't know how to weigh the interests of future generations against the current generation, then we don't yet know

<sup>1</sup> For the relevant estimates, see GiveWell, 'Against Malaria Foundation', November 2016, <http://www.givewell.org/charities/against-malaria-foundation/November-2016-version>.

how we ought to act in response to climate change. If we don't know how to weigh the interests of distant strangers against compatriots, then we don't yet know the extent of our duties to the global poor. We aren't going to resolve these difficult moral questions any time soon. But we still need to act now. So we need to know how to act, despite our uncertainty.

Given the prevalence and importance of moral uncertainty, one would expect ethicists to have devoted considerable research effort to the topic of how one ought to make decisions in the face of moral uncertainty. But this topic has been neglected. In modern times, only one book and fewer than twenty published articles deal with the topic at length.<sup>2</sup> The book you are reading attempts to begin to address this gap.

In this book, we address the questions of whether there are norms that are distinct from first-order moral norms that govern how one ought to act given one's fundamental moral uncertainty and, if so, what those norms are.

These questions raise many difficult theoretical issues, and we don't pretend to have comprehensive solutions to all of them. Our aim, instead, is to offer an up-to-date introduction to the topic, make a first pass at solving some of these issues, and to invite others to build on this work. Though we cover many topics, the core of our argument is to defend an information-sensitive approach to decision-making under moral uncertainty: accepting that different moral views provide different amounts of information regarding our reasons for action, and that the correct account of decision-making under moral uncertainty is sensitive to that. Ultimately, the default account we defend is a form of *maximizing expected choiceworthiness*. We defend various departures from this default position for cases in which expectation is not well-defined.

Before we begin, let us clarify some terms and delimit the scope of this book. When we refer to 'moral uncertainty', we use 'moral' in the broad sense, referring to uncertainty about what we all-things-considered morally ought to do. We can distinguish this from the even broader term 'normative uncertainty', which also applies to uncertainty about which

<sup>2</sup> We say 'modern times' because there was also extensive discussion of similar issues by Catholic theologians, such as Bartholomew Medina, Blaise Pascal, and Alphonsus Liguori. See Bartolomé de Medina, *Expositio in primam secundae angelici doctoris D. Thomae Aquinatis*, 1577; Blaise Pascal, *Lettres Provinciales*, 1657; Alphonsus Liguori, *Theologia Moralis*, 2nd edn, 1755. For a summary of this discussion, see F. J. Connell, 'Probabilism', in Thomas Carson (ed.), *The New Catholic Encyclopedia*, 2nd edn, Detroit, MI: Thomson/Gale, 2002. For discussion of this debate and its relevance to the modern debate on moral uncertainty, see Andrew Sepielli, "Along an Imperfectly Lighted Path": Practical Rationality and Normative Uncertainty', PhD thesis, Rutgers University, 2010, pp. 46–51.

theory of rational choice is correct,<sup>3</sup> and uncertainty about which theory of epistemology is correct. A full treatment of these additional issues would warrant a book of its own, so we have chosen to focus exclusively on moral uncertainty.

There are also issues relevant to moral uncertainty that, for reasons of focus, we do not address, except briefly. We do not thoroughly address the issue of whether moral ignorance is exculpatory in the same way that empirical ignorance is exculpatory, though we discuss this briefly in Chapter 1.<sup>4</sup> We do not significantly discuss the extent to which one should alter one's moral beliefs in light of moral disagreement, and, apart from a short discussion in the first chapter arguing that we ought to be at least reasonably unsure in our moral views, we do not discuss the question of what credences one ought to have in first-order moral theories. Finally, simply to remain focused, we do not attempt any significant discussion of the long-running debate within Catholic theology about what to do when different Church Fathers disagreed on some moral matter.<sup>5</sup>

Instead, the focus of this book is firmly on the question:

Given that we are morally uncertain, how ought we to act in light of that uncertainty?

We make the following structural assumptions about what a decision under moral uncertainty looks like. We consider a decision-maker choosing from a set of jointly exhaustive and mutually exclusive options ( $A, B, C, \dots$ ). These options could be acts, or they could be plans of action, or anything else that could be the subject of choice and moral assessment.

We suppose that the decision-maker has credence in each of a set of first-order moral theories ( $T_1, T_2, T_3, \dots$ ). We will normally talk about these theories as if they are complete stand-alone moral theories, such as a particular form of utilitarianism. However, they could often just as well represent partially specified theories, or particular moral considerations regarding the options at hand, such as whether killing is equivalent to letting die.

We will sometimes represent the credence in a given theory with a real number between zero and one. This is not to assume that we have precise

<sup>3</sup> For an introduction to this issue, see William MacAskill, 'Smokers, Psychos, and Decision-Theoretic Uncertainty', *The Journal of Philosophy*, vol. 113, no. 9 (September 2016), pp. 1–21.

<sup>4</sup> For a discussion of this issue, see Elizabeth Harman, 'Does Moral Ignorance Exculpate?', *Ratio*, vol. 24, no. 4 (December 2011), pp. 443–68.

<sup>5</sup> See footnote 2 for references to some of the literature on this topic.

credences in these theories. Nothing will turn upon the exact values of these credences, and we believe that everything we say could just as well be said if we were to use imprecise credences. In this book, we remain agnostic on whether theories of moral uncertainty should be specified with respect to decision-makers' actual credences or to their epistemic credences (that is, the credences they ought, epistemically, to have). Everything we say could apply given either choice.<sup>6</sup>

We will assume that the theories under consideration assess these options in terms of *choiceworthiness*, which represents the strength of the reasons for choosing an option. This need not be quantitative: it could just provide an ordering of which options are more choiceworthy than others. We will often consider theories which can make at least roughly quantitative judgments about choiceworthiness, such that one option might be slightly more choiceworthy than a second, but much more choiceworthy than a third. We will occasionally use numbers to represent these levels, and define a *choiceworthiness function* as a numerical representation of a theory's choiceworthiness ordering such that a higher number represents a more choiceworthy option. Apart from the subsection on supererogation in section IV of Chapter 2, where we discuss the issue of the relationship between choiceworthiness and deontic status, we'll call an option *permissible* (right) iff it is maximally choiceworthy (that is, iff there is no option that is more choiceworthy than it in the option set), and *impermissible* (wrong) if it is not maximally choiceworthy. Occasionally, where it is more natural to do so, we'll talk about 'severity of wrongness' or 'moral value' rather than choiceworthiness; we mean this to refer to the same concept.

Some decisions made under moral uncertainty are intuitively superior to others. For example, intuitively there is something important to be said against choosing option *A* when all theories in which you have credence consider it to be impermissible, and they all consider option *B* to be permissible—even if, according to the true moral theory, action *A* is the morally correct choice. We shall use the term *appropriate* to make such assessments of options under moral uncertainty, where *A* is more appropriate than *B* iff a rational and morally conscientious agent who had the same set of options and beliefs would prefer *A* to *B*.<sup>7</sup> As we use the term, to say that an act is

<sup>6</sup> For an argument that the theory should be specified in terms of epistemic credences, see Andrew Sepielli, 'How Moral Uncertainty Can Be Both True and Interesting,' *Oxford Studies in Normative Ethics*, vol. 7 (2017), pp. 98–116. <https://www.oxfordscholarship.com/view/10.1093/oso/9780198808930.001.0001/oso-9780198808930-chapter-6>

<sup>7</sup> We put aside cases where this account of appropriateness will give the wrong results, such as when a decision-maker who is not in fact morally conscientious faces a situation where an

appropriate is to say that no alternative option is more appropriate than it. More than one option can be appropriate, some options may be more appropriate than others, some may be *incomparable* in appropriateness, and there may be degrees of appropriateness.

This framework allows us to more precisely state the central question of this book:

For any given set of credences in moral theories and set of options that a decision-maker can have, what is the appropriateness ordering of the options within that option set?

We shall generally assume descriptive certainty, though it is of course possible to simultaneously have descriptive and moral uncertainty. This is just to simplify things.<sup>8</sup>

Our approach to answering the central question is as follows. We look at the different *informational situations* that decision-makers can find themselves in with respect to the theories they face, where an informational situation is determined by the way in which choiceworthiness can be compared both within each theory in which the decision-maker has credence, and across those theories.

In this book, the approach we take is ‘divide and conquer’. We ask, for each of a number of different informational situations, what the correct theory of decision-making under moral uncertainty is given that informational situation. As an analogy for this approach: one might argue that, under empirical uncertainty, maximize expected value is the correct theory when one has determinate credences across all possible outcomes, but that maximin is the correct theory when one has no clue what credence to assign to different outcomes.

There is a wide range of possible informational situations, and in this book we will not be able to go through them all. We hope to demonstrate

evil demon has set things up such that a certain action is good only if the decision-maker is morally conscientious. Accommodating cases like this is not important for the project of this book.

<sup>8</sup> For an interesting argument that one cannot plausibly take both moral and empirical uncertainty into account at the same time, see Ittay Nissan-Rozen, ‘Against Moral Hedging’, *Economics and Philosophy*, vol. 31, no. 3 (November 2015), pp. 349–69. However, we don’t find his argument compelling. If you are motivated to take moral uncertainty seriously, and you are genuinely unsure about how risk-averse you ought, morally, to be, then you should not find what Nissan-Rozen calls ‘Standard Dominance’ plausible; and if you don’t find it plausible, then his argument has no bite.

Table 0.1

		Comparability Conditions			
		Full comparability	Unit-comparability	Level-comparability	Incomparability
Measurability Conditions	Ratio-scale	✓	✓	✗	✗
	Interval-scale	✓	✓	✗	✓
	Ordinal scale			✗	✓
	Preorder			✗	✓

the fruitfulness of the divide and conquer approach; we do not pretend to be comprehensive in its application. We therefore lay out the main possible informational situations in Table 0.1. We indicate with a tick which set-ups we consider, and shade out those informational conditions that are not possible:

The measurability of a theory describes which *intratheoretic* comparisons of choiceworthiness can be made, where the different measurability conditions we highlight are as follows.

First, a theory can give a preorder. If so, then the choiceworthiness relation is transitive (for all  $A, B, C$ , if  $A$  is at least as choiceworthy as  $B$ , and  $B$  is at least as choiceworthy as  $C$ , then  $A$  is at least as choiceworthy as  $C$ ), and reflexive (for all  $A$ ,  $A$  is at least as choiceworthy as  $A$ ), but it is not complete (where completeness is the property that for all  $A, B$ , either  $A$  is at least as choiceworthy as  $B$  or vice-versa.) We therefore cannot represent the theory with a choiceworthiness function. A choiceworthiness preorder would naturally result from a theory on which some values are incomparable.

Second, a theory can give *ordinal scale measurable* choiceworthiness. On such theories, the choiceworthiness relation is transitive, reflexive and complete (therefore ranking options as 1st, 2nd, 3rd (etc.) in terms of choiceworthiness) and the relation can therefore be represented with a choiceworthiness function. However, such theories don't give any information about *how much* more choiceworthy the most choiceworthy option is, rather than the second most choiceworthy. More precisely: Let  $CW_i$  be a numerical representation of  $T_i$ 's choiceworthiness ordering, such that

$CW_i(A) > CW_i(B)$  iff  $A$  is more choiceworthy than  $B$  on  $T_i$ . If  $T_i$  is ordinal scale measurable, then  $CW_j$  also represents  $T_i$  iff  $CW_j = f(CW_i)$ , where  $f(x)$  is any strictly increasing transformation.

Third, theories that provide *interval-scale measurable* choiceworthiness give us not just ordinal information about choiceworthiness, but also tell us the ratio of differences in choiceworthiness between options. More precisely: If  $T_i$  gives interval-scale measurable choiceworthiness and  $CW_i$  is a numerical representation of  $T_i$ 's choiceworthiness ordering, then  $CW_j$  also represents  $T_i$  iff  $CW_j = kCW_i + c$ , where  $k$  and  $c$  are real numbers with  $k > 0$ .

Fourth, theories could also potentially provide *ratio-scale* measurable choiceworthiness, in which case they would have a non-arbitrary zero point, and give meaning to ratios between the absolute levels of choiceworthiness of options. More precisely: If  $T_i$  gives ratio-scale measurable choiceworthiness and  $CW_i$  is a numerical representation of  $T_i$ 's choiceworthiness ordering, then  $CW_j$  also represents  $T_i$  iff  $CW_j = kCW_i$ , where  $k > 0$ .

The comparability of two or more theories describes which *intertheoretic* comparisons of choiceworthiness can be made, where the different comparability conditions we highlight are as follows.

If two moral theories are *unit-comparable*, then we can meaningfully make claims about the ratio of differences in choiceworthiness between options across theories: we can say that the difference in choiceworthiness between  $A$  and  $B$  on  $T_i$  (where  $A$  is more choiceworthy than  $B$  on  $T_i$ ) is  $k$  times as great as the difference in choiceworthiness between  $C$  and  $D$  on  $T_j$  (where  $C$  is more choiceworthy than  $D$  on  $T_j$ ).

Whether or not they are unit-comparable, two theories might also be *level-comparable*. If two theories are level-comparable, then we can meaningfully say that the choiceworthiness of one option, on one theory, is greater than, equal to, or less than, the choiceworthiness of another option on the other theory.

If two moral theories are *fully comparable*, then the *intertheoretic* comparisons of choiceworthiness that can be made between theories are the same as the *intratheoretic* comparisons of choiceworthiness that can be made within each theory. So, for example, two interval-scale measurable theories that are fully comparable are both unit-comparable and level-comparable; two ratio-scale measurable theories that are fully comparable are both level-comparable and ratio-comparable (where we can compare the ratios of levels of choiceworthiness across both theories).

If two moral theories are *incomparable*, then they are neither unit- nor level-comparable. We cannot say that the difference in choiceworthiness

between two options on one theory is larger, smaller or equally as great as the difference in choiceworthiness between two options on the other theory; nor can we say that the level of choiceworthiness of one option on one theory is greater, smaller, or equal to the level of choiceworthiness on the other theory.

We believe that Table 0.1 provides at least the primary informational situations of interest. But this table could be expanded. Though we doubt that such an idea is meaningful, one could potentially consider theories on which choiceworthiness is measured on an absolute scale (where no transformation of the theory's choiceworthiness function is permissible). More interestingly, one could also consider situations of ratio-scale or interval-scale measurability with intratheoretic incomparability; the meaningfulness of such a notion has been shown by Erik Carlson.<sup>9</sup>

Within those informational situations that we have listed above, we are able to investigate in depth only three, which we regard as particularly important: interval-scale measurability with unit-comparability, interval-scale measurability without unit or level-comparability, and ordinal scale measurability without level-comparability. Because we don't discuss conditions of level-comparability in this book, when we refer to *intertheoretic comparability* we are referring in every instance to unit-comparability.

We restrict ourselves to these informational conditions just to make things easier for ourselves: this is only the second modern book written on the topic of decision-making under moral uncertainty and we have to pick our battles if we are to make progress at all. However, in Chapter 6, we do briefly discuss how our account might be able to handle theories with incomplete choiceworthiness orderings. We do hope, though cannot argue here, that many other informational conditions can be treated in a similar way to how we treat the informational conditions we do consider.<sup>10</sup>

Of the informational situations that we don't discuss, one has been studied by other philosophers: Christian Tarnsey and Ron Aboodi have

<sup>9</sup> Erik Carlson, 'Extensive Measurement with Incomparability'. *Journal of Mathematical Psychology*, vol. 52, no. 4 (2008), pp. 250–9. We also note the possibility of multidimensional scales, and different scales to account for various infinite number systems (such as the extended reals, transfinite ordinals, infinite cardinals, hyperreals and surreals; we thank Christian Tarnsey for emphasizing this.

<sup>10</sup> For example, Christian Tarnsey ('Rationality and Moral Risk', dissertation, pp. 181–2) argues that binary structure (where a theory simply puts all options into two categories, 'permissible' and 'impermissible' and says nothing more) is importantly distinct from ordinal structure. Whether or not that is true, we are inclined to treat the two informational conditions in the same way, using the Borda method to aggregate both forms of uncertainty.

discussed what to do in conditions of ordinal measurability and level-comparability.<sup>11</sup> They argue in favour of stochastic dominance as a condition of adequacy on any theory of decision-making under such conditions. We find their approach promising—though there is much more work to be done in order to develop a complete theory—but simply for reasons of focus we are not able to discuss their work in this book. We don't know of work that addresses the other informational conditions. We believe that studying these informational conditions is a ripe area for further work.

As we will see, a decision-maker under moral uncertainty can face more than one of these informational situations at one and the same time, when theories in which the decision-maker has credence differ in how we can make choiceworthiness comparisons within or between them. We discuss this in Chapter 4.

With this terminology and these clarifications in hand, we can describe the structure of the book, as follows. In Chapter 1, we introduce the topic of moral uncertainty and argue that we should take moral uncertainty seriously, in particular arguing that there is a meaningful sense of 'ought' that is relative to moral uncertainty.

In Chapter 2, we will show that the problem of moral uncertainty cannot be solved by just saying that we should follow the moral theory we have most credence in, or by just saying that we should choose the option that is most likely to be morally right. Instead, we argue that one should treat empirical and normative uncertainty analogously and that, therefore, what we should do in cases of moral uncertainty depends upon both the decision-maker's credences over moral theories and the degrees of choiceworthiness that those theories assign to options. More specifically, we argue that, in conditions where all the moral views in which we have credence are both interval-scale measurable and intertheoretically comparable and we have well-defined credences, we should *maximize expected choiceworthiness*. We defend this idea against two objections: that the account is too demanding, and that it can't account for theories that allow for supererogation.

In Chapters 3–5, we discuss what we consider to be the most serious problems facing any account similar to *maximize expected choiceworthiness*: that sometimes choiceworthiness is not comparable across different moral

<sup>11</sup> Christian Tarsney, 'Moral Uncertainty for Deontologists', *Ethical Theory and Moral Practice*, (forthcoming); Ron Aboudi, 'Is There Still Room for Intertheoretic Choiceworthiness Comparisons?', MS, University of Toronto.

theories and that sometimes theories do not even give meaningful quantities of choiceworthiness. In Chapter 3, we introduce an analogy between decision-making under moral uncertainty and the problem of social choice, and show how this allows us to develop principles for decision-making under moral uncertainty even when faced with theories that provide merely ordinal choiceworthiness and are non-comparable. In Chapter 4, we extend this work to address the situation where a decision-maker is faced with theories that do give meaningful quantities of choiceworthiness but are not comparable with each other, and then propose a general account of decision-making under moral uncertainty, which can be viewed as an extension of *maximize expected choiceworthiness*. In Chapter 5, we discuss the question of when, if ever, moral theories are comparable with each other, arguing against some accounts of intertheoretic comparisons that have been proposed in the literature and sketching our own novel account.

In Chapter 6, we discuss two key problems for any account of decision-making under moral uncertainty: the problems of fanaticism and infectious incomparability. We argue that the information-sensitive account defended in previous chapters allows us to give a satisfactory solution to these problems.

In Chapters 7–9, we discuss certain metaethical and practical implications of the idea that one ought to take moral uncertainty into account in one's decision-making. In Chapter 7, we discuss the apparent conflict between moral uncertainty and non-cognitivism, arguing that the existence of moral uncertainty poses a significant problem for non-cognitivists. In Chapter 8, we examine the implications of moral uncertainty for debates in practical ethics, and argue that in the literature so far the application of moral uncertainty to practical ethics has been simplistic. In Chapter 9, we introduce the concept of the value of moral information, and show how this has implications for the value of engaging in ethical reflection and study.

Let us now turn to our substantive arguments. We begin by arguing that we should take moral uncertainty seriously, and that our central question is a non-trivial one.