Moral Uncertainty and Permissibility: Evaluating Option Sets[*]

Christian Barry
School of Philosophy, Research School of the Social Sciences, Australian National University, Canberra, Australia

Patrick Tomlin
Politics and International Relations, University of Reading, Reading, UK

Abstract: In this essay, we explore an issue of moral uncertainty: what we are permitted to do when we are unsure about which moral principles are correct. We develop a novel approach to this issue that incorporates important insights from previous work on moral uncertainty, while avoiding some of the difficulties that beset existing alternative approaches. Our approach is based on evaluating and choosing between option *sets* rather than particular conduct *options*. We show how our approach is particularly well-suited to address this issue of moral uncertainty with respect to agents that have credence in moral theories that are not fully consequentialist.

Key words: moral uncertainty; Ted Lockhart; consequentialism and non-consequentialism; moral-belief-relative permissibility.

I Introduction

With respect to issues that are the subject of heated moral and political debate, regarding them as uncertain or even suspending belief about their moral status often seems to be the most sensible positions to take.[1] However, whatever we may decide to believe about these issues (if we decide to believe something about them at all), we will often find ourselves in situations where we must conduct ourselves in ways that reflect practical attitudes toward them.

---

[1] Throughout this essay we will use the term 'uncertain' in the colloquial sense—the condition of being in doubt—rather than in the technical sense that is commonly employed in decision theory—circumstances in which one cannot assign probabilities to various outcomes.

To take a commonplace example, suppose that Sue could either send her son to an excellent private school, or send him to an adequate state-funded school, enabling her to donate the money she would have spent to charities tackling global poverty. Let us imagine that Sue is uncertain whether it is permissible to send her son to the private school, given what else she could do with the money. However uncertain she may be about this matter, she either will or will not send her son to the private school. If she does send him, then she acts in a way that seems inconsistent with a sincere commitment to the view that it is impermissible to send him to the school. We then must face the issue, as Andrew Sepielli (2009) has aptly described it, of what to do when we don't know what to do. Given how ubiquitous uncertainty about the ethical status of different choices is, it is surely a matter of major philosophical and practical importance how we should behave, and how we should decide how to behave, when we are morally uncertain. This issue has, however, only recently begun to receive systematic attention from moral philosophers.[2]

In this essay we seek to advance debate on moral uncertainty by proposing a distinctive approach to these issues. This approach advises uncertain agents to consider and evaluate different *option sets* rather than considering each option on its own. So far as we know, all philosophers who have, to date, proposed solutions to the problem of moral uncertainty have focused on individual options. Doing so, we argue, fails to take into account key non-consequentialist concerns, and so will fail to address any form of moral uncertainty in which non-consequentialist theories play a role. In contrast, our approach can handle a mixture of consequentialist and non-consequentialist theories, as well as situations where only consequentialist theories are in play.

---

[2] Some noteworthy recent contributions to this nascent literature are: Ross 2006; Moller, 2011; Jackson and Smith 2006; Guerrero 2007; Lockhart 2000; Sepielli 2006, 2013a, 2013b; Weatherson 2002.. For an overview, see Sepielli (forthcoming).

We used the term 'approach' rather than, say, 'theory' advisedly. This is because we do not pretend to have settled all the issues here. In particular, the difficulty of inter-theoretic comparisons of value, something which plagues most approaches to the problem of moral uncertainty, is amplified on our view (we will describe below how this is so). In addition, there may be several different variants of our approach which nevertheless agree with our fundamental contention – that option sets, not individual options, should be the unit of comparison. In our view, the nascent debate on moral uncertainty has gone in the wrong direction, and we hope to offer an alternative path here. We hope that in laying out that alternative path, we will inspire others to come with us, and help us to develop it more fully.

II Preliminaries

In terms of conduct, the general question posed by moral uncertainty is this: given that I am unsure about what I should do, what should I do? It might be thought that the answer to this question is obvious: I should do the morally right, or perhaps morally best, thing. This is not wrong, since of course I should, objectively speaking[3], do the morally right, or perhaps morally best, thing – I ought to act in accordance with morality. But I need to know how I ought to act, given what is known to me at the time when I must decide how to act. And information about whether one or another course of conduct is the morally right, or best, option is not the sort of thing that is known to me, since by hypothesis I am torn between rival views about what morality

---

[3] For the purposes of this essay, we will not make any assumptions about the status of ethical claims. We will sometimes speak of moral claims or principles being 'objective', 'right,' 'correct' or 'true', but could rephrase these claims in terms more friendly to anti-realists (e.g., as the moral claims or principles to which the agent should commit herself). On moral uncertainty and non-cognitivism, see Sepielli 2012. We will also speak of options being 'morally impermissible' and will assume that such a prohibition indicates an all-things-considered prohibition. However, we do not think the plausibility of anything we say here turns on this assumption.

demands and/or recommends. I need guidance about what I should do, *given* that I am unsure about what morality demands and/or recommends.[4]

To evaluate the substantive merits of different approaches to what we ought to do under conditions of moral uncertainty, it is important to distinguish two importantly different questions that we can ask concerning what we should do in the face of such uncertainty. Consider the following case.

Pregnancy

Sue must decide whether or not to terminate her pregnancy.

Sue is unsure whether or not she should terminate her pregnancy. It would be in her self-interest to be without the unwanted foetus, but she wants to act morally. She might ask two distinct moral questions: 'given my uncertainty, what would be the *morally best* or *most admirable* thing for me to do?' and 'given my uncertainty, what is *permissible*?' (i.e., 'what options may I *legitimately choose from*?'). In this paper we restrict ourselves to this latter question concerning permissibility under moral uncertainty.

Now, of course, many philosophers, at the first-order level, believe the questions of moral permissibility and moral optimality to have the same answers, since they believe all morally sub-optimal options to be impermissible. This does not mean, however, that they fail to recognise that the two questions are distinct; and so it is not morally controversial, nor does it assume that the answers to the questions differ, to separate the questions.

---

[4] It is important to note here that we focus exclusively on what we should choose to do after we have deliberated. We assume that (first-order) deliberation has taken place, and has resulted in uncertainty. We say nothing about the appropriate attitude towards our uncertainty, or our epistemic duties when uncertain. We are grateful to Nic Southwood for discussion of this point.

The same is true at the second-order level of moral uncertainty. In separating the questions of permissibility and moral optimality under conditions of moral uncertainty, we do not assume that the answers to these questions are different. But we do allow that they may be. When it comes to dealing with questions of moral uncertainty, it is those who *fail* to separate the questions of what, when uncertain, is morally best and morally permissible who make an unwarranted controversial assumption, since they cannot properly accommodate first-order theories that not only recognise the existence of two different questions, but acknowledge that the answers to them can differ. Given this, and that the difference between the questions of permissibility and optimality are well-recognised at the first-order level, it is perhaps surprising that, so far as we know, the two questions have not been fully separated in the nascent moral uncertainty literature to date.[5]

Why focus on permissibility? In first-order moral assessment, we recognise a fundamental distinction between claims that conduct is impermissible and claims that conduct is morally sub-optimal. Judgements of impermissibility carry a special kind of moral gravity, and impermissible actions attract a special kind of moral blame. Therefore, it seems appropriate to treat the issue of permissibility under uncertainty as particularly pressing.

Philosophers have worried about the *nature* of the oughts that theories of moral uncertainty try to capture.[6] It could either be a rational or a moral ought. In our view (which we cannot fully defend here) moral uncertainty theories deliver a (certain kind of) moral ought. However, it should be noted that we do not think that whether

---

[5] See, for example, Lockhart 2000, ch. 5, where he argues against the concept of 'supererogatory' actions, and proceeds to address moral rightness and moral bestness together. While Lockhart's arguments may be good ones, this remains too controversial an assumption to form part of a theory of moral uncertainty, or at least one that hopes to govern the decisions of anyone who has credence in theories which allow for supererogatory actions.

[6] See, in particular, Weatherson 2002. We are grateful to referees of this journal for encouraging us to be more explicit on our understanding of the nature of the ought.

or not someone accepts our *substantive* approach to moral uncertainty will necessarily depend upon whether they accept our understanding of the nature of the 'ought'. (Of course, if they deny that there is any sense in talking about what we ought to do, and what is permissible, under moral uncertainty, they will reject our specific view about what is permissible.)

We think that the ought in question is a moral ought because we believe that those who act contrary to their moral beliefs commit a moral error. Imagine George and Fred both perform the same (objectively) impermissible action. George knew it to be impermissible, while Fred (reasonably) had no idea. Both have committed a moral wrong. But it seems to us that George commits an additional *moral* wrong to Fred – he does wrong, but he also does what he *knows* to be wrong. Now imagine George didn't *know* the conduct was impermissible, but strongly suspected it to be so (and had many other options to choose from). Again, it seems that this would be a moral error. Given this, we are inclined to view the verdicts of moral uncertainty theories as moral directives.[7]

This may seem to present us with a worry: how can morality deliver different verdicts - the objective verdict and the uncertainty verdict? We do not think that this is a major concern. As Derek Parfit (2011, 150-164) has persuasively argued, we should accept a plurality of senses of 'ought' (and cognate terms such as 'permissible'). Just as we should accept 'fact-relative', 'evidence-relative', and 'belief-relative' senses of ought in relation to empirical facts, evidence, and belief, so

---

[7] It may be that we ought to think of Fred as excused but that George is not. This may well be correct, but this then raises the question: under what conditions can moral uncertainty or error excuse? Our paper can be read as an answer to that question. We are grateful to a referee for useful comments here.

we should accept objective, evidence-relative, and belief-relative moral oughts in relation to moral facts, evidence, and belief.[8]

In essence, then, what we are investigating here is the question of *moral-belief-relative permissibility*. The idea of this kind of permissibility is that it defines what is permissible *given our moral beliefs*. Of course, our moral beliefs may be false (or at least some of them may be), and we may even be blameworthy for holding some of them. Nevertheless, given our beliefs, it seems like moral-belief-relative permissibility is the relevant standard for deciding which options are permissible under conditions of moral uncertainty. This can be contrasted with objective permissibility, the kind of permissibility that philosophers are usually interested in.

In the next few sections we will present, assess and reject a variety of approaches to moral-belief-relative permissibility under moral uncertainty. In doing so, we will show how these approaches give direction to agents that, we argue, is implausible. Note that when we criticize approaches to decision-making under moral uncertainty because they allow or disallow some course of conduct, we do not mean that it seems implausible that a view treats some course of conduct as objectively morally permissible or impermissible. *That* type of claim is the type to be considered in first-order moral deliberation. In judging whether a theory of moral decision-making under conditions of moral uncertainty is plausible, we must instead ask whether it is plausible that, *given this agent's beliefs*, the conduct in question is moral-belief-relative (im)permissible.

---

[8] In fact, these two sets of distinctions, along empirical and moral lines, will cross-cut. So it will make sense to ask whether some conduct is empirical-fact- and moral-belief-relative permissible. Or whether some conduct is morally objectively and empirical-evidence-relative permissible. To avoid these complications, we will assume throughout that the all the empirical facts are known.

III The My Favourite Theory Approach

One way to approach the issue of moral uncertainty is for the agent to weigh up the best arguments for different moral theories or principles, and then be guided only by the theory or principle she finds most plausible. This is, tacitly, how most moral philosophers appear to approach the problem—they offer arguments for a course (or courses) of conduct they deem to be permissible or best, and hope to convince others. Following Ted Lockhart (2000, 42-43), we will refer to this as the My Favourite Theory Approach to moral uncertainty. Once an agent has done her best to decide which conduct options are permissible or morally best, lingering doubts about whether or not she is really confident that these conduct options have this moral status are treated as irrelevant—she should act on these options as if she were not in doubt about their moral status.

If we are unsure about some matters of empirical fact that pertain to the question of how we ought to act, that should influence how we act. We don't act as if something is *definitely true* just because we know it's *possible* that it is true, or even *likely* to be true. And yet the My Favourite Theory Approach would have us ignore uncertainty about moral principles when it comes to acting. Consider **Pregnancy**.[9] Imagine that Sue is 51% confident that abortion is permissible, but accepts that if it *is* impermissible then it would be a serious moral wrong, akin to murder, to have an abortion. She is sure that carrying the foetus to term is (in her case, at least) permissible. The My Favourite Theory Approach would render abortion moral-belief-relative permissible, even though she is 49% confident that it is objectively very seriously wrong. On this view she should view herself as moral-belief-relative

---

[9] For a detailed discussion of this particular type of case see Moller 2011.

permitted to abort the foetus even though she believes it is nearly as likely that she will be committing murder as that she will be acting objectively permissibly.

We regard the advice given by the My Favourite Theory Approach in this case (and others like it) to be highly implausible. In **Pregnancy** Sue should not (given her beliefs) regard herself as moral-belief-relative permitted to have an abortion, and consequently she ought not to have an abortion. We wouldn't ordinarily perform some action that carried a 49% chance (or, for that matter, a 20% chance) of being a serious moral wrong akin to murder (unless moral reasons concerning the alternatives rendered that the best of a bad bunch of options), so why should the fact that the uncertainty in question is moral uncertainty change that? We think that our uncertainty in these cases should properly be reflected in what we are moral-belief-relative permitted to do.

IV Sophisticated Theory Selection

It seems important that the fact that we are morally uncertain be reflected in how we decide to act – an agent take into account *all* the principles or theories in which she credence when deciding what to do, not just those views in which she has most credence.

But how should we incorporate our uncertainty? Crucially, in tackling the question of moral-belief-relative permissibility under moral uncertainty, we must decide whether we should be seeking a moral theory to act in accordance with or a particular course or courses of conduct. In other words, should our theory of permissibility under moral uncertainty seek to directly pick out one (or more) of the conduct options available to us, or should it seek to pick out one of the theories we hold credence in for us to adopt for the purposes of action?

The My Favourite Theory Approach recommends looking for a theory, and it has a very simple rule about how to obtain the right theory to act on —that is, we should employ the one we find most plausible. We might, however, seek out a more sophisticated way of finding a moral theory to act on. Such an approach might be inspired by[10] Jacob Ross's (2006, 743) suggestion that we should 'accept' (i.e., use for action-guidance)[11] a theory based on the following considerations:

> in deciding what theory to accept, although we have *pro tanto* reason to prefer more probable theories to less probable ones, we also have *pro tanto* reason to prefer theories according to which the differences or dispersion among the values of our options is higher to theories according to which this dispersion is lower. We may call this the 'dispersion principle'. (Ross 2006, 758).

On Ross's view, how much credence we place in a moral theory is an important factor in choosing what theory to accept; but so too is what is at stake, and there is more at stake in theories with higher 'dispersion.' Accordingly, we need to measure how much is at stake (Ross suggests the difference between the best option and the average value of the options) and multiply this by the (subjective) probability that the theory is true.[12] This will give us the expected value of accepting a given theory. We might be tempted to apply this more sophisticated way of selecting a theory, or something like it, to our question concerning moral-belief-relative permissibility, and accept and act in accordance with the theory with the highest expected value.

---

[10] Ross's (2006) theory, whilst the inspiration for this putative approach, differs in that (a) Ross's question is not quite the same as ours (it doesn't pick out permissibility as the sole focus); (b) Ross's theory is about 'acceptance' and he does allow that there could be degrees of acceptances (743-744), and therefore does not, at least at the outset of the article, demand that we accept only one theory (although at times he does seem to be addressing the question of 'What theory should we accept?' (760)). However, Ross does not tell us how to move from partial acceptance to action-guidance. Given that the drive behind the distinction between belief and acceptance is practical, this is surprising, and leaves the issue of what we are to do unanswered. We are grateful to a referee for useful comments here.

[11] In Ross' terminology, 'to accept a theory is to aim to choose whatever option this theory would recommend' (2006, 743)

[12] Ibid., p. 759.

However, any approach which focuses on selecting *a theory* to follow under uncertainty is vulnerable to damaging objections. However sophisticated the choice mechanism, in asking us to choose *a* theory to act on, such an approach gets things wrong at the first step. Consider an example, **Emergency,** in which Sue must decide what she may permissibly do when confronted with ten drowning people. She has three conduct options but is unsure which of two theories is correct. Imagine that the conduct options available to Sue are the following:

Option A is killing one person to save ten;

Option B is mildly harming one person to save eight;

Option C is leaving the bystander alone and allowing the drowning people to die.

Theory 1 is a form of consequentialism, while Theory 2 gives significant weight to distinctions such as doing and allowing harm, and intending and foreseeing harm. On Theory 1 it is *seriously* wrong to allow ten people to drown, and permissible (indeed, required) to kill an innocent bystander to save them. It is impermissible to harm someone to save eight, as it does not bring about the best consequences (saving ten), but more wrongful to do nothing. On Theory 2, it is *seriously* wrong to kill an innocent, less wrongful to harm an innocent and permissible (indeed, required) to leave the bystander alone. Sue assigns a 'zero' score to those options that are permitted and then a negative score to impermissible actions. The worse an impermissible option is, the lower the score. The payoffs associated with the conduct

options (according to the theories in which Sue has some credence) are represented in

Table 1.[13]

Table 1

	Theory 1	Theory 2
Option A	0	-20
Option B	-2	-2
Option C	-20	0

We suggest that, if she places roughly equal credence in both theories, Sue ought to

regard Option B as moral-belief-relative permissible, indeed required, even though

both theories in which she places credence would condemn Option B as objectively

impermissible. Under any form of theory selection, however, she would be required to

regard, as moral-belief-relative impermissible something (namely, Option B) that,

intuitively, she should regard as what she ought to do, given her commitments.

　　Let's suppose that Sue thinks that it is slightly more likely that Theory 1 is the

correct theory. As such, she assigns a 51% probability to Theory 1 and 49% to Theory

2. My Favourite Theory would tell the agent to act as if Theory 1 were definitely true,

and so perform Option A. This seems implausible, since it risks a likely serious

wrong, when Option B—viewed as a much less serious wrong by both theories—is

---

[13] We assume here, for the purposes of argument, that comparisons of the wrongness of conduct are possible *across* moral theories. Such an assumption is controversial. It is not, in fact, necessary for the point in question, as theory selection is implausible even when only ordinal rankings of options are available—imagine an option is selected as second best across all theories, while each theory's best option is ranked very low by all other candidate theories. It seems that the only option that would be ruled out by theory selection—the universally acknowledged second best—is the one we should perform. The literature on moral uncertainty has begun to deal with the problem of inter-theoretic comparisons, but it is not a topic we take up in this paper. For interesting discussion of the problem and important steps towards solving it, see Ross 2006, 761–775; Sepielli 2006; 2009; 2013a. It seems worthwhile to us to see what kinds of principles are desirable *if* such comparisons are possible. This way, we can see which kinds of comparison it would, in fact, be useful to be able to make, and direct our attentions accordingly. We will return to this point in Section IX.

available. Sophisticated Theory Selection would also require her to adopt Option A. Because each of the two theories is equally 'dispersed,' Sue should accept, and act on, the theory that she views as most likely to be correct.

If we are right that Option B is the right action to perform under moral uncertainty, then a seemingly paradoxical conclusion arises. For Sue, Option B is moral-belief-relative permissible, *even though she does not believe it to be permissible.*[14] While this may *seem* paradoxical, it is not. That is because two different kinds of permissibility are in play. She should regard B as moral-belief-relative permissible, and as objectively impermissible. Importantly, moral-belief-relative permissibility does not only include those options we believe to be objectively permissible, and includes some that we do not. This is not as surprising as it may seem. After all, factual-belief-relative permissibility can include options that we know are fact-relative impermissible.[15]

V The Hierarchical Approach

Rather than any form of Theory Selection, we ought instead to develop a view which does not seek to provide guidance based on one particular theory or another. Instead, our moral-belief-relative permissible options should be constructed anew, rather than seeking to endorse the options provided by a particular theory. One approach of this type is considered by Ted Lockhart in his pioneering work *Moral Uncertainty and its Consequences* (2000). Lockhart(Lockhart 2000, 26) considers the following normative principle as a guide to rational (or reasonable, he uses these terms interchangeably) action under conditions of moral uncertainty:

---

[14] These comments were prompted by helpful comments from a referee for this journal.
[15] See, for example, Parfit's (2011, 159) mineshaft case. In that case, because of uncertainty about the facts, we should choose to do something (allow ten to drown) that we know is fact-relative impermissible (since it would be avoidable if we knew all the facts).

PR2

In situations of moral uncertainty, I (the decision-maker) should (rationally) choose some action that has the maximum probability of being morally [permissible].[16]

That is, we ought to opt for the option that has the lowest risk, according to all the principles we are entertaining, of being impermissible.

To illustrate, consider, for example, how this approach would guide Sue in **Pregnancy**. As Lockhart (2000, 51-52) notes, while the question of whether it is morally permissible for a woman to abort her foetus is a very morally contentious and difficult one, it is almost always considered morally permissible for the woman to carry the foetus to term and give birth. If Sue is certain that carrying to term is permissible and there is a chance that terminating the foetus is impermissible, PR2 would maintain that Sue ought to regard herself as permitted (indeed, required) to carry to term, since not having the abortion carries *no* risk of being impermissible whilst having the abortion carries *some risk* of being impermissible. A PR2-inspired answer to the question of moral-belief-relative permissibility under principle uncertainty is the Hierarchical Approach (HA), so called since it establishes a lexical hierarchy among conduct options in terms of the moral status that the deliberating agent supposes them to have:

1. Conduct that is certainly permissible;

---

[16] Lockhart's statement of PR2 refers to 'morally right' rather than 'morally permissible' conduct. One reason we have substituted the term 'permissible' is to keep the terminology consistent. Another is that it is hard to know whether we are on the same ground as Lockhart here. Lockhart (2000, 5) makes it clear that he equates 'rightness' with 'permissibility'. However, since he believes that there is conceptually no difference between moral bestness and moral rightness (ch. 5), and that there is no firm line between rightness and wrongness (and thus, permissibility and impermissibility) (ch. 4), we are not confident that Lockhart's theory addresses the same question as we do. If this turns out to be the case, we are happy to simply say that we are adapting Lockhart's insights to attempt to answer our (different) question.

2. Conduct that may be permissible;

3. Conduct that is certainly impermissible.

HA

Any option that falls within category 1 must always be chosen prior to all options in category 2, and any option in category 2 must always be chosen prior to all options in category 3.

At first glance, the HA may seem intuitively plausible. If a conduct option is certainly impermissible, we are morally forbidden from performing it. Why, then, would we prefer such an option over one that may not be forbidden? It is prima facie plausible to think that the very *first* thing that a theory of conduct under moral uncertainty must do is to remove those options *we know* to be impermissible from the table. And if we must choose between an act that we know to be permissible and an act that may be impermissible, should we not obviously choose the former? Nevertheless we believe that *neither* the lexical priority of options in category 2 over options in category 3 *nor* of options in category 1 over those in category 2 is to be found in the best theory of permissibility under moral uncertainty.

Regarding the priority of possibly permissible options over certainly impermissible options, this becomes less plausible when we consider that not all impermissible options are equal—some are worse than others. We have already shown, above, that in cases such as **Emergency**, we should perform certainly impermissible options (i.e., Option B) that are not serious moral wrongs over possibly permissible options that may also be very serious moral wrongs (i.e., Options A and

C). This leads us to the surprising yet compelling conclusion that moral-belief-relative permissibility can include options that we know to be objectively impermissible.

VI Minimizing Expected Impermissibility

We can reject the priority of possibly permissible options over certainly impermissible options if we adopt an Expected Moral Value (EMV) approach. (This is in fact what Lockhart does.[17]) The EMV approach tells us, through judging the 'moral score' of a conduct option according to a theory, multiplying it by our credence in that theory and then summing the weighted scores of each option, to pursue the option or options with the highest expected moral score. When we are looking at moral-belief-relative *permissibility* for conduct, then it would seem appropriate to Minimize Expected Impermissibility (MEI).[18]

> **MEI**
>
> Under moral uncertainty, the moral-belief-relative permissible option(s) are those which minimize expected impermissibility, (where 'expected

---

[17] When Lockhart (2000, 82) moves beyond describing conduct options as simply 'right' and 'wrong' he posits this principle, which exemplifies the EMV approach:
> PR4: In situations in which moral agents are uncertain of the degrees of moral rightness of some of the alternatives under consideration, a choice of action is rational if and only if the action's expected degree of moral rightness is at least as great as that of any other alternative.
As previously stated (see n. 16), although Lockhart explicitly equates 'rightness' with 'permissibility', he has other commitments which make it unclear whether we are addressing the same question. A similar approach appears to be advocated in Sepielli 2009 although it is, again, not obvious that Sepielli is answering the 'permissibility' question with which we are concerned.

[18] This is essentially to create an asymmetry between permissible and impermissible options. Permissibility/impermissibility is a binary distinction, but not all impermissible actions are on a moral par (e.g., promise-breaking as opposed to killing) – call this *degree of wrongfulness* – and neither are all permissible actions – call this *degree of praiseworthiness*. MEI takes into account expected wrongfulness but not expected praiseworthiness, i.e., it takes into account how far below the permissibility line an option (potentially) falls, but not how far above. We think this is correct. If one thinks two options have a 50% chance of being permissible, but if impermissible Option A is akin to murder and Option B is akin to promise-breaking, only Option B should be seen as moral-belief-relative permissible – i.e., expected wrongfulness matters. On the other hand, if one is certain that two options are permissible, but one has a higher expected praiseworthiness, this will make no difference whatsoever to moral-belief-relative permissibility – i.e., expected praiseworthiness is irrelevant (to the question we are interested in here).

impermissibility' is a function of likelihood of impermissibility and severity of wrongfulness).

This would mean that category 2 options would no longer enjoy lexical priority over category 3 options, but that category 1 options would continue to enjoy lexical priority over category 2 options. Since category 1 options are certain to be permissible, then if one such option is on the table it will be preferred to any possibly permissible option, since there is no risk of wrongdoing if one adopts it.

The absolute priority of certainly permissible actions has significant intuitive appeal, and many of Lockhart's arguments support it forcefully. But there are problems with this approach. Consider cases in which an agent has some credence in two (or more) moral theories where one theory 'nests' inside the other. 'Nesting' occurs when one theory (the more permissive theory) regards as permissible all the options deemed permissible by the other (more restrictive) theory, as well as some additional options. To give a concrete example, in **Pregnancy** the position that carrying to term is the only morally permissible option 'nests inside' the position that *both* abortion and carrying to term are morally permissible.

In such cases, HA and the related MEI approach can be far too demanding. If you demand that when conduct that is certainly permissible is available only certainly permissible options are moral-belief-relative permissible, this implies that for the purposes of deciding what to do we must act in accordance with the most restrictive theory that we have *any* credence in. And it requires that we do this even if we have very little credence in this most restrictive theory, and even if the potential wrongdoing would be slight, according to this theory. Consider **Giving**.

Giving

Sue is a person with a relatively good standard of living who has just received $30,000 and is deciding how to spend it. She must decide whether and to what extent she should contribute to poverty relief.

Sue is 99% convinced of a morality according to which the money is rightly hers, and that the amount she may give to charity is wholly at her discretion. It may be morally better for her to give away a substantial portion of her newfound gains, but she does not act impermissibly if she decides instead to spend it on herself or her near and dear. However, Sue has recently come across Peter Singer's work, according to which, 'when we spend our surplus on concerts or fashionable shoes, on fine dining and good wines, or on holidays in faraway lands, we are doing something wrong' (Singer 2009, 19). Although not *entirely* convinced by Singer's arguments, Sue has been swayed to some extent, so that she now believes that there is a very small chance (1%) that Singer is correct, and that the only permissible course of action is to give away all of her gains. However, she does not believe that she does anything morally *dreadful* if she chooses to keep the money – that is, even if keeping the money is wrong, she does not consider it to be a serious wrong.

The lexical priority of certainly permissible options over possibly permissible options requires us to believe that Sue is moral-belief-relative permitted only to give away all of her money, since this is the only option that both first-order theories in which she has credence regard as permissible. Thus, even though she thinks that keeping any of the money represents a very small chance of a very small (objective) wrong, she should give it all away. Remember, what we are asking ourselves is '*given Sue's beliefs*, what is she permitted to do?'. In this case, Sue has little credence in

Singer's principle, yet that theory is the one that HA and MEI would have her adopt for the purposes of action. We believe the lexical priority of certainly permissible over possibly permissible options is too demanding *given what Sue believes*.

Now consider a case of what we call the 'Venn' variety. Like nesting cases, in Venn cases there are some options which both the moral theories under consideration regard as permissible, and therefore those options are certainly permissible. However, each theory regards some further options to be permissible which the other regards as impermissible. To take a simple example (which we will call **Suicide** for reasons that will soon become clear), imagine that Sue must decide whether to visit a friend or a lonely stranger in hospital. Theory 3 says visiting the friend and not the stranger is impermissible. Theory 4 says that visiting the stranger and not the friend is impermissible. Neither theory regards these actions as *seriously* wrong, though. Both theories, however, place central importance on Sue's control over the end of her own life, and as such affirm that suicide is permissible.

According to HA and MEI, Sue is moral-belief-relative *required* to kill herself (as that is the only certainly permissible option), in order to avoid an (objective) minor wrong (visiting the friend or the stranger). This result seems especially troubling if Sue regards the claim that you are objectively morally *required* to kill yourself as morally repugnant, and so would reject as repugnant any theory of objective morality which claimed this. The first-order theories Sue places credence in *allow* suicide because they consider it of central importance that Sue is morally permitted to decide when and how her life ends (so far as is possible). Yet on HA and MEI Sue finds herself moral-belief-relative *required* to commit suicide. Her first-order moral beliefs about the importance of being morally permitted to decide

somehow lead to a moral-belief-relative instruction which allows her no such decision.

Many of us are committed to the view that suicide is morally permissible. But we are also morally uncertain about a great many things. Given this uncertainty, the chance of our going through life without committing some objective moral wrong – even if we always act in accordance with what we believe to be permissible, or try to minimise moral impermissibility – is certainly greater than zero. That is, by living we, *at the very least*, run a small risk of committing a minor wrong. If HA or MEI are correct, many of us should kill ourselves just to avoid this small risk.

This Venn case is importantly different from the nesting cases like **Giving.** In those cases, under HA and MEI what is moral-belief-relative required (e.g., giving away the money) is something that she at least believes *may* be objectively morally required. In the Venn case, the overlapping area that becomes moral-belief-relative required, because it is the only option that is certainly permissible (e.g., Sue killing herself), is something she is sure is not objectively required.

That you may be moral-belief-relative required to do something that you are sure is not objectively required is counter-intuitive. But we do not think it is a fatal flaw for a theory of moral uncertainty. After all, in **Emergency** we endorsed viewing Option B as moral-belief-relative required *even though Sue is sure it is (objectively) impermissible*. But there is an important difference between these two cases. In **Emergency**, Sue's discomfort was with Option B *as an option*, and she was able to register that discomfort through her scoring of the options. In **Suicide**, because she views suicide as certainly permissible, any theory of moral uncertainty that focuses only on whether options are permissible or impermissible (and, if so, how seriously)

renders her unable to register *any* discomfort with, let alone extreme aversion to, the instruction 'suicide is required'.

Lockhart (2000, 108-110) recognises some of the difficulties with MEI that we have identified. However, he tries to deny that cases like **Giving** are in fact nesting cases, since against demanding Singer-like theories we will hold credence *not* in theories that allow us to do with our money what we will, but rather theories that will regard giving away all our money as impermissible due to duties we have to ourselves and our nearest and dearest. Thus, Lockhart implies, **Giving** cases will actually be what we shall call 'cross-condemning' cases, in which those actions permitted by one theory are condemned by the other, and vice versa (e.g., **Emergency**). We are not convinced by this response. Many people will *not* hold credence in such theories which regard giving away all of our money as impermissible. So, the HA must still deal with its counterintuitive implications for those agents for whom **Giving** is indeed a nesting case. In addition, even if theories which include duties to ourselves and those with whom we enjoy special relationships condemn Sue for giving away *all* her money, their influence on the final decision as to what she should regard as permissible will be minimal. Most theories that require (and not only allow) her to prioritise herself and/or her nearest and dearest will surely nevertheless allow her to give away the vast majority of the $30,000 (especially as we imagine that Sue is already comfortably well-off before this windfall).

In order to defend the priority of the certainly permissible, Lockhart uses the example of uncertainty regarding abortion. Many readers may be unsympathetic to a position that so quickly entails that we should regard abortion as moral-belief-relative

impermissible,[19] but it is hard to resist the conclusion that if an agent genuinely believes that there is a chance that by having an abortion she will do something morally akin to murder, she should not run this risk when a certainly permissible option is on the table (see Moller 2011). However, note that there are two special features of the abortion case which make assigning lexical priority to certainly permissible acts over possibly permissible acts appear more plausible. These are the sharp binary nature of the choice situation—an agent either can abort or not, there are no intermediate options—and the magnitude of the potential wrong involved—the taking of the life of an innocent being with full moral status.[20] Affirming the lexical priority of certainly permissible options over possibly permissible options will seem much less plausible when these two features are not present, and one need not affirm this priority to conclude that in this case Sue should not have the abortion. For example in **Giving** there are many options between giving nothing and giving everything, and wrongdoing that would be risked in failing to give all of it is (by the agent's own lights) minor. Our belief is that focusing on this case, we can develop an intuitive response which would require rejecting both MEI and HA.

Let's say that in **Giving**, Singer's theory requires Sue to give $30,000. Suppose also that a competing theory in which she has credence says that she is entitled to give whatever she chooses (in other words, $0 is permissible and so is $30,000). She is uncertain about which of the two theories is correct. For reasons we have discussed above, both the HA and MEI state that giving $30,000 is moral-belief-relative required, for this option is permitted by both candidate theories, whilst giving $0 is permitted by only one theory.

---

[19] It must be noted here that even if one believes that abortion is *morally* impermissible, or should be regarded as moral-belief-relative impermissible, it does *not* follow that it should be made illegal.

[20] That is, the wrong would be great (we assume) *according to the moral theory that she holds credence in that condemns abortion*. Moral theories that condemn abortions usually condemn them as serious wrongs akin to killing the innocent.

But there are other options: the minimum amount that Sue is moral-belief-relative permitted to donate is not limited to being either $0 or $30,000. Indeed, in this situation it seems most reasonable that Sue should regard herself as required to give away *some* of the money (reflecting her belief in the Singer theory) and as permitted to do what she wants with the rest (reflecting her much stronger belief in the non-Singer theory).[21]

VIII Moral Uncertainty and the Consequentialism/Non-Consequentialism Debate

Much of the appeal of regarding certainly permissible options over possibly permissible options (and thus HA and MEI) appears to rely on the idea that no theory 'loses out' when her moral-belief-relative permissible options are restricted to the options that all the theories endorse as permissible. Since all the options endorsed by a more restrictive theory as permissible are also regarded as permissible by the less restrictive theory, an agent choosing a conduct option endorsed by the more restrictive theory is a 'win-win'—the option will be regarded as permissible by both theories. As such, that option is seemingly Pareto superior to any alternative conduct option that is viewed as *im*permissible by any potentially true theory. Since the question we're considering is 'what conduct option(s) are moral-belief-relative *permissible*?' then the relevant 'data' appear to be what each candidate theory regards as objectively permissible, and from this perspective, acting on the more restrictive theory appears to be endorsed by both theories.

---

[21] This includes the option of giving it all away.

However, it seems to us that HA and MEI fail to take into account a potentially[22] important additional moral consideration—how the competing moral theories view the *range or combination of options* that are morally permitted, and *why* they recommend that range of options as permissible. Consider **Giving**. In this case, the theory that says Sue's money is hers to do whatever she chooses with (in which Sue places 99% credence) says that both giving and not giving are permissible, while Singer's theory (1%) says that only giving it all is permissible. The thought behind HA appears to be 'since nothing is lost, in terms of permissibility, by giving the $30,000, she should view herself as required to give it. There is no objection from the first, non-Singer, theory (of which she is 99% convinced) to her giving the $30,000.'

There is an important difference, however, between her being *permitted* to, and choosing to, give $30,000 and her being *required* to give $30,000. While the non-Singer theory is indeed indifferent between the actions of giving nothing and giving $30,000, it is *not* indifferent between her having the option set of giving $0–$30,000 as opposed to an option set that consists only of giving $30,000.

Similarly, in **Suicide**, the theories that Sue places credence in regard it as important that she have full moral discretion over whether she lives or dies. In terms of permissibility, these theories are indifferent to whether she decides to kill herself or continue living. But they are *not* indifferent between her being *required* to kill herself, versus merely being permitted.

In many moral theories (particularly those that incorporate non-consequentialist elements), giving people a range of morally permitted options matters a great deal. This means that options do not matter only as individual options; the way that options are combined also matters. The permissible option *set* matters,

---

[22] That is, it may be viewed as important by some of the moral theories she places credence in, not that we ourselves are undecided about its importance.

not just the individual options. This kind of concern is reflected in complaints that some moral theories (and the option sets they endorse) are too demanding or too restrictive. These are complaints about the option sets recommended by certain moral theories.

HA and MEI are strongly biased against theories that value a range of permitted options, or specific option *sets*, rather than specific options themselves, as they only allow theories to register complaints (as it were) to the extent to which other theories allow *specific options* which they condemn. They do not allow theories to register complaints about option sets being unduly limited, or take account of the *reasons* why theories regard options (or groups of options) as permissible. We need an approach that is better placed to properly represent the concerns of such theories under moral uncertainty. Since moral-belief-relative permissibility needs to take account of our moral beliefs, when our moral beliefs are not only concerned with individual options, neither should our theory of moral-belief-relative permissibility.

We think that this problem may have been missed in the literature thus far because those who write on the problem of moral uncertainty often appear to adopt a broadly consequentialist perspective. This is perhaps not surprising. Consequentialists have simple and compelling approaches to empirical risk and uncertainty, and it is quite tempting to apply those approaches to moral uncertainty. But since consequentialists tend to see the best option as the only permissible option, and thus the required option, they tend not to see much distinction between what is permissible and what is required. Indeed, when consequentialists recognise a plurality of options as permissible, it is because they all just happen to maximise the good equally well. Therefore, the consequentialist does not seek to defend the permissibility of an option as part of a broader set of options, while at the same time considering the idea that it

is the only required option to be morally repugnant – either an option maximises the good or it doesn't. But this isn't how non-consequentialists think. They are concerned to provide agents with a *range of protected options*[23], and so are concerned with the permissible option set as a whole. One can think suicide rightly forms part of the permissible option set, whilst regarding a theory which requires suicide as repugnant.

To be clear, on non-consequentialist theories, this range of protected options should not be concerned only with absence of interference. When non-consequentialists worry about demandingness or restricted choices, they are not worried only about other agents forcing us to do what is demanding or restricting our choices. They are concerned about the demandingness and restrictiveness of morality itself.

Approaches to permissibility under moral uncertainty, like HA and MEI, that focus only on the moral 'score' of individual options do not appear to be able to take into account such concerns, and as such, they cannot properly mediate between moral theories when some of the theories in question incorporate or are founded on these concerns.

MEI may be the best decision mechanism when we are unsure between two consequentialist theories that are indifferent to the range or combination of options available to the agent. Therefore, one lesson to draw here could be that the appropriate theory of permissibility under principle uncertainty may depend on the range of first-order moral theories in which the agent has some credence. When theories share certain assumptions or positions concerning morality or its structure, those assumptions can be taken for granted in trying to mediate between them. But when the moral theories disagree on some issue, any decision procedure that neglects

---

[23] For Frances Kamm this is one of the central tenets of non-consequentialism. For an overview of the approaches one make take in defending such a position, see Kamm 1992.

entirely the core features of one or more of them will not provide a plausible approach to moral-belief-relative permissibility under those conditions.

However, we are not yet prepared to give up the search for a *general* theory of permissibility under moral uncertainty. We believe that a new approach – Evaluating Option Sets – can be applied both when only consequentialist theories are present and when non-consequentialist theories enter the fray.

VIII An Alternative: Evaluating Option Sets

Our analysis thus far suggests that the HA and MEI approaches are unable to properly take account of moral theories which place importance on agents having *option sets* and thus may find the *requirement* to do something that they regard as *permissible* as morally troubling. In deciding what is moral-belief-relative permissible, HA and MEI look to the surface only, seeing where the theories we place credence in 'end up'— which *specific options* they endorse. We think that, if a theory of moral uncertainty is to properly respect theories that care about option sets, and their restrictiveness and demandingness, a new approach is required.

As a result, we propose the Evaluating Option Sets approach. This approach is formed of the following five principles:

Evaluating Option Sets

1. A theory of moral-belief-relative permissibility should seek to identify a *set of options* that are permissible (though this set may include only one option)

2. The potential option sets to be evaluated should not be limited to those endorsed by the moral theories in which the agent has credence, but rather should be the entire possible range of options sets.

3. Each option set must be evaluated from 'within' each first-order theory. This is done by evaluating, from the perspective of the first-order moral theory:

    a. The individual options it contains (looking at whether they are permissible or impermissible, and if impermissible, how seriously wrongful they would be); and

    b. The option set as a whole. In scoring the option set as a whole, a theory should give its own preferred option set the highest score (a perfect 0, if option sets are 'complained about' through negative scores).[24]

4. Each option set should then be given an overall 'score'. In order to do this, the scores generated by Principles 3a and 3b should be combined. They could be summed. This provides the overall score for an option set from 'within' a theory. This score is then multiplied by the (subjective) probability that the theory is correct. This is then repeated for each theory, and the scores from within each theory are summed to provide the option set's overall score.

5. The option set with the highest score wins. All options within the winning option set are moral-belief-relative permissible.

As we indicated at the outset, we regard this as a promising, but currently incomplete, approach. In particular, there will be complications and difficulties within principles 3

---

[24] We make this stipulation for two reasons. First, it makes the comparison between option sets (viewed as option sets) a comparative exercise: a theory should 'complain' (through negative scores) about option sets which it views as *unduly* restrictive. A theory should always 'back its own horse' by giving its preferred option set the highest score, and uniquely so – otherwise on what basis can it claim that the option set is the best one? Second, without this stipulation, allowing theories to give their own option set less-than-perfect scores can result in counter-intuitive outcomes. We are grateful to a referee of this journal for pointing out this danger.

and 4 of the above statement, and different variants of the Evaluating Option Sets approach may attempt to solve them in different ways. In what follows, we will seek to outline the view in general, through an example, and also to explain how we think the kinds of calculations required by principles 3 and 4 should be undertaken. We will also point toward some difficulties that any variant will face.

As an example as to how the Evaluating Option Sets approach works, imagine a choice situation in which there are three options – A, B and C. Imagine the agent has credence in two theories. One says that only A is permissible. The other says that all of option set (ABC) are permissible. (This is, therefore, a nesting case, like **Giving**).

Evaluating Options Sets tells us to look for an option set to endorse under moral uncertainty (Principle 1). But we are not to choose merely between those option sets endorsed by the first order moral theories ((A) and (ABC)). Instead, we look at the whole possible range (Principle 2). Here, there are seven potential option sets: (ABC), (AB), (AC), (BC), (A), (B), (C). Each must be evaluated.

How do we evaluate option sets? First, we must evaluate them according to each first order theory (Principle 3). How is this done? Well, option sets are not, of course, completely independent of their contents, and thus a large part (and for some theories, the whole) of ranking the option set will be a matter of ranking its constituent parts in terms of expected impermissibility. Suppose Theory 1 (a consequentialist theory), says that only A is permissible (and thus required), that B is mildly impermissible (-1) and C is seriously impermissible (-20). We now must turn these various evaluations into an overall assessment of the option set's individual options (Principle 3a). One way to do this would be to sum the total, such that the

option set scored (from within Theory 1) -21 (0-1-20). We think this is mistaken.[25] This is because of the following kind of case. Imagine an option set which contains one thousand options, all of which are impermissible but not serious wrongs. They are each scored at -1. Now consider an option set that contains only one option, a very serious wrong, scored at -1000. On the summing approach, these two options would be viewed as equals (all else equal). But surely this is an error. After all, the agent can only perform one of the options (in any given moral decision) and so she can either perform one of the very minor wrongs, or the one very serious wrong. These are not equivalent choices. Instead, we propose to evaluate the options within the option set with an average figure. Since each option (if the option set is the winner) will be equally permissible at the bar of moral-belief-relative permissibility, each should be viewed as equally choice-worthy. And since the agent can choose any option, but can only choose one, we should look at the average expected impermissibility of the set. Thus, (ABC) should score -7 (-21/3) from within Theory 1.

That figure will tell us how the option set should be viewed (from within Theory 1) qua its individual options. However, on this approach the evaluation of the individual options is not (necessarily) all there is to evaluating option sets. For those theories that take it to be important that a *range* of options are morally permissible, the evaluation of an option set will *also* incorporate an evaluation of the option set *qua option set* (Principle 3b). This is the key innovation of our approach. It allows moral theories to score the option sets *qua option sets* that its competitors recommend (as well as some option sets that no theories recommend). So, from the perspective of Theory 1, the agent is able to evaluate the option set recommended by Theory 2, not

---

[25] We came to realise this due to helpful comments from a referee.

just in terms of its individual options (Principle 3a), but as a *set*, allowing the theory to register complaints about the demandingness or restrictiveness of other option sets.

For example, take a non-consequentialist theory (Theory 2) in which it is thought important to provide an agent with a range of options because this respects persons as ends in themselves, rather taking them to be mere producers of the good (see, for example, Kamm 1992, 358-359). Imagine this theory takes A, B *and* C to be permissible. In this case, all seven option sets would score a perfect 0 when looked at as individual options (Principle 3a), since they all contain only permissible options. But, different option sets may or may not manage to treat the person as an end in herself, or do so to a greater or lesser degree. Imagine that the option set (AB), for example, scores particularly well (on Principle 3b), achieving a -2 overall, but (A) is too restrictive, failing to treat the agent as an end in herself, and is given -10. (ABC), as Theory 2's preferred option set, should receive a 0.

Our consequentialist theory (Theory 1) does not care for option sets qua option sets. It cares about maximizing the good, not the way in which individual options interact with each other. From within Theory 1, we think that each option set (when evaluated qua option set – Principle 3b) should receive a perfect 0 – no option set should be marked 'down' for its failure to provide something that the theory does not care for. The theory cares about the options, and it is in the first stage (Principle 3a) that it is given the opportunity to register 'complaints' about these. It may be, however, that this way of handling the option set as a whole is controversial. This is one place where different variants of the Evaluation Option Sets approach may differ.

Once each option set has been scored, both qua individual options, and qua option set, from each first-order theory, we can calculate their overall scores (Principle 4). Imagine that the agent places a 40% credence in Theory 1, and a 60%

credence in Theory 2. Now take option set (A). From Theory 1, it scores a 0 both qua options and option set. From Theory 2, it scores 0 for options, but -10 for option set. These two scores must then be summed (again, this is possibly a controversial way to get from the scores generated at Principles 3a and 3b to an overall assessment from a first-order theory). To get the overall figure for an option set, taking all theories into account, we must multiply the scores from the first-order theories by the subjective probability that the theory is true. Therefore, overall, option set (A) would score -6.[26]

Recall that this is a nesting case. On the prevailing approaches to moral uncertainty, option set (A) would win, since A is the only certainly permissible option. On our view, however, (AB) will beat it. That is because while (AB) would lose out from the perspective of Theory 1 because it contains a minor wrong, it would beat A from the perspective of Theory 2 as an option set overall. (AB) scores -1.4 overall.[27] (AB) would also beat (ABC), as that would be dragged down by the presence of C, which would score very badly from the perspective of Theory 1. Therefore, (AB) would be the preferred option set, and both A and B would be moral-belief-relative permissible options.

This approach coheres with our intuitive judgments about a range of cases. For example, a requirement to commit suicide, even when suicide is regarded as certainly permissible, would, when the agent views required suicide as repugnant because of the failure of such an option set to deliver what it is she cares about when she endorses *permissible* suicide (namely choice about the end of one's life), score very badly compared with the minor wrong of visiting the wrong person in hospital. But the (potential) serious wrong of killing an innocent being of moral standing could

---

[26] (0x0.4)+(-10x0.6)

[27] (AB) from Theory 1: -0.5 for options + 0 for option set = -0.5 overall.
(AB) from Theory 2: 0 for options + -2 for option set = -2 overall
(AB) overall = (-0.5x0.4)+(-2x0.6) = -1.4

mean that in **Abortion,** viewing oneself as moral-belief-relative required to carry to term may be the appropriate stance to take, even considering the limited option range. In **Giving**, given that Sue has very little credence in the Singer theory, and she thinks the potential wrong minor, a range of options should be viewed by her as moral-belief-relative permissible, although this range will not include withholding all of the money.

IX Objections and Problems

In this final section, we will outline some objections and difficulties with our view as stated. Some we feel able to answer. Others we can simply raise and, at best, give a first response.

The first objection is that by focusing on option sets, Evaluating Option Sets adopts a non-consequentialist view of morality and as such incorporates a controversial moral view. We do not think that our approach incorporates elements of, or is unduly biased toward, non-consequentialist theories in any damaging sense. It tries to take into account the concerns of some non-consequentialist theories, to be sure, but those concerns will influence the outcome of any decision-making under moral uncertainty only to the extent that the agent *already* holds these concerns at the level of first-order morality. Our theory will provide the same outcomes as MEI when the 'inputs' are two (or more) theories that place no weight on the range of options open to agents. In such cases, an option set's score will be determined purely by the scores of its constituent options, which will lead to the selection of the option set which minimizes expected impermissibility.

A second, and related, worry is that this approach 'double counts' concerns over demandingness and the range of options available to agents. We think this is

mistaken. It is true that such concerns are what drove the agent to place credence in the permissibility of a given option set (e.g., (ABC)) in the first place. But under MEI that concern does not affect what she is moral-belief-relative permitted to do. Under MEI, she, in effect, submits a list of discrete options, and is nowhere able to register discomfort with or preferences for how those options interact. As an analogy, consider two professors who must select which students to give a prize to. Professor 1 submits the name of one student, A. Professor 2 has been very impressed by some group work done by students A, B, and C. He wishes to award the prize to them jointly. However, Professor 2 does not think that any individual in this group should be awarded the prize on their own, and would prefer that no prize be awarded rather than it be given to any one of these individuals. Under a system in which each professor simply submits a list of names, and those names are taken as discrete recommendations, there is no way for Professor 2's belief in the equal deservingness, and the *joint* deservingness, of the candidates, nor his belief that no award is preferable to a single award, to be taken into account. Anyone who looked at the lists, saw A on both, and awarded the prize to A for this reason, would not have taken Professor 2's views adequately into account.

Our belief is that MEI is in a similar position. While a concern to avoid over-demandingness and/or treat agents as ends in themselves may drive the non-consequentialist theory's list of preferred options, if the theory's entry into the uncertainty decision-making procedure can only take the form of a list of options, its concerns regarding option sets, demandingness, and restrictiveness cannot be taken into account.

A third problem concerns inter-theoretic comparisons of value. As we observed above (see introduction, and n. 13), one of the most difficult problems

regarding theories concerning moral uncertainty is the issue of inter-theoretic comparisons of value. In order for Theory 1's judgment of Action C being -20 to be commensurate with Theory 2's judgment of it as 0, the two theories must be trading in the same currency. Our theory, unfortunately, further complicates this already difficult issue. First, it requires options to be measured in terms, not of their expected 'value' in the sense of 'goodness' (which is how a purely consequentialist approach taking consideration of only consequentialist theories might do things), but rather their wrongfulness. This means that we need consequentialist theories to convert value into permissibility/impermissibility, and then that must be converted back into a value – the wrongfulness. However, we think that on all plausible moral theories, impermissible options can nevertheless be better and worse. Killing is worse than promise-breaking on the non-consequentialist account, and doing no good is worse than doing some, but not enough, good on the consequentialist account. Indeed, our approach may be on comparatively comfortable ground here. Insofar as each candidate theory is trying to provide judgments about moral permissibility, it is at least trading in the same currency as all the others. This is not true of opposing consequentialist theories which have different accounts of the good.

Second, in order to apply our approach there must be, at the least, *two* kinds of value that can be compared across theories: the values of options (as mentioned above, and as would also need to be the case for MEI and Sophisticated Theory Selection) and the values of option sets. Third, these two kinds of value must, ideally, be commensurable, or at least comparable, with each other as well.[28] We do not know if this is possible. It seems to us that we can provide these valuations intuitively. It is, after all, a single agent who must provide the various values: the values may be inter-

---

[28] This needn't necessarily be the case however – each option set might attract two different scores, one based on options and the other on the set as a whole, and the selection of the final option set may be based on intuitively weighting the two values.

theoretical, but they are intra-agent. But our task here is not to try to solve these knotty issues: our approach is to try to inform the debate about inter-theoretic comparisons of value through identifying the best theory of moral-belief-relative permissibility under moral uncertainty. In doing so we can see which kinds of comparison we would ideally like to be able to undertake. We believe that the Evaluating Option Sets theory is a plausible and original answer to the normative question, and as such should inform work on the difficult issue of comparisons that plagues all theories of moral uncertainty.

The final problem concerns option sets that receive equal scores.[29] To see why this is a problem, consider a choice situation in which the agent faces three options: D, E, and F. Now consider two theories, which score the options as follows:

	Theory x	Theory y
D	0	0
E	0	-4
F	-4	0

Now, imagine that the agent places equal credence in the two theories. This would mean, looking solely at individual options, the option sets would score as follows: (D) = 0; (DE) = -1; (DF) = -1; (DEF) = -1.33

Imagine (D) scores very poorly on the overall option set calculation, and so is eliminated from consideration, while (DE), (DF) and (DEF) all score equally on that measure. We now face the following puzzling situation: (DE) and (DF) score equally. What should we tell the agent? The obvious thing is to tell her to choose between these two option sets. But if we do this, she effectively has a choice of options D, E

---

and F. But (DEF) is itself a defeated, and inferior, option set. We are unsure how best to deal with this problem. We think that perhaps she should choose an option set, and then stick to it. This isn't *quite* the same as having (D,E,F) as her option set, since it is a two stage process (choose option set, choose option). Practically-speaking, however, it seems near identical.

X Concluding Remarks

Evaluating Option Sets allows agents to carve out a range of legitimate options in a way that respects all of the moral concerns that they believe may apply to the situation (with regard to permissibility), and the credence they place in various theories (and thus those concerns). Other theories seem to have focussed simply on 'outputs' (individual options), rather than looking to what range of options different moral theories allow, and *why*. We believe, therefore, that Evaluating Option Sets offers a better approach to the question of moral-belief-relative permissibility. There are some issues with the approach, but we believe it is worth pursuing until such issues are shown to be insurmountable. The most difficult issue the theory faces – inter-theoretic comparisons of value – is one that besets all theories.

Of course, in emergency situations and the like, it is unlikely that we will take out our pen and paper (or smart phone) and perform the kinds of calculations that the theory demands. But this is an issue that all approaches to moral uncertainty face. In our view we should get the view right 'on paper' first. This, in turn, may lead to the development of easier to use heuristics and rules of thumb.

REFERENCES

Guerrero, A.A. 2007, "Don't Know, Don't Kill: Moral Ignorance, Culpability and Caution." *Philosophical Studies* 136: 59–97.

Jackson, F. and Smith, M. 2006. "Absolutist Moral Theories and Uncertainty." *Journal of Philosophy* 103: 267–283.

Kamm, F.M. 1992. "Non-Consequentialism, the Person as an End-in-Itself, and the Significance of Status." *Philosophy & Public Affairs* 21: 354-389.

Lockhart, T. 2000. *Moral Uncertainty and its Consequences*. Oxford: Oxford University Press.

Moller, D. 2011. "Abortion and Moral Risk." *Philosophy* 86: 425-443.

Parfit, D. 2011. *On What Matters*, Vol, 1. Oxford: Oxford University Press.

Ross, J. 2006. "Rejecting Ethical Deflationism." *Ethics* 116: 742–768.

Sepielli, A. 2006. "Review: Ted Lockhart, *Moral Uncertainty and its Consequences*." *Ethics* 116: 601–604.

Sepielli, A. 2009. "What to Do When You Don't Know What to Do." In *Oxford Studies in Metaethics*, vol. 4. Edited by R. Schafer-Landau. Oxford: Oxford University Press.

Sepielli, A. 2012. "Normative Uncertainty for Non-Cognitivists." *Philosophical Studies* 160: 191-207.

Sepielli, A. 2013a. "Moral Uncertainty and the Principle of Equity among Moral Theories." *Philosophy and Phenomenological Research* 86: 580-589.

Sepielli, A. (2013b). "What to Do When You Don't Know What to Do When You Don't Know What to Do..." *Nous* 47: 521-544.

Sepielli, A. (forthcoming). "Moral Uncertainty." In *Routledge Encyclopedia of Philosophy Online*.

Singer, P. 2009. *The Life You Can Save*. Melbourne: Text Publishing.

Weatherson, B. 2002. "Review: Ted Lockhart, *Moral Uncertainty and its Consequences*." *Mind* 111: 693–696.