Moral Uncertainty and the Criminal Law[*]

Christian Barry (ANU) and Patrick Tomlin (Warwick)

Forthcoming in The Palgrave Handbook of Applied Ethics and the Criminal Law
Edited by Kimberly Kessler Ferzan and Larry Alexander

I Introduction

Our aim in this chapter is not to connect an *issue* in applied ethics to the criminal law, but rather to take a nascent, evolving *approach* to applied ethics, and to see what it might have to say about the kinds of questions and positions we find within criminal law theory.

The approach we want to apply to the criminal law is that of what we will call 'Moral Uncertainty Theory'. Applied ethics, broadly speaking, concerns what, morally speaking, real life agents ought to do in particular real life circumstances. Ordinarily applied ethicists argue for the position that they believe to be the moral truth of the matter.[1] So, for example, Philosopher A writes an article arguing that the correct moral position is that abortion is morally permissible, while Philosopher B writes an article arguing that the correct moral position is that abortion is morally impermissible.[2]

This is all well and good, and of course a lot of fine philosophical work gets done in this way. But it's also important to remember the *applied* nature of applied ethics – it is supposed to offer guidance to real life agents. Imagine that Jane is unsure about whether or not to have an abortion. She thinks it's probably permissible but is deeply unsure. She reads the work of Philosopher A and Philosopher B. This only heightens her sense of uncertainty. She gets the chance to speak to Philosopher A. Philosopher A tries to convince Jane that the moral truth is that she is permitted to have an abortion. But then Philosopher B tries to convince her of the opposite. Jane remains deeply uncertain about what the right course of action is. What is Jane to do?

This is where Moral Uncertainty Theory comes in. Moral Uncertainty Theory tries to advise Jane on what she ought to do *given her uncertainty*. It starts from the idea that Jane is uncertain, and asks what she ought to do given the moral beliefs that she has. Moral Uncertainty Theory asks the question that Jane is surely asking herself: *what should I do, given that I don't know what I should do?*

It might be thought that the answer to this question is obvious: Jane should respond by doing the objectively morally right thing (Weatherson 2014; Harman 2015). Well, of course she should! But Jane needs to know how she ought to act, given what she believes at the time when she must decide how to act. And she is uncertain about whether or not some of the courses of action open to her are permissible. Perhaps there is another obvious answer to our question: Jane should suspend belief about the matter. That is, when we cannot tell with much confidence whether or not some course of conduct is permissible, we can simply suspend belief about it.

However, whatever Jane may decide to *believe* about competing claims regarding the permissibility of abortion, she is in a situation where she will conduct herself in ways that *reflect a practical stance* towards this practice. Suppose that Jane is committed to the view that, if a foetus possesses the same moral status as a toddler, she has a stringent responsibility not to terminate her pregnancy. Even if she suspends belief about the claims presented by A and B, she must adopt a practical stance towards the issue of abortion. For example, if she terminates her pregnancy, then given her beliefs, she is taking the practical stance that a foetus does not possess the same moral status as a toddler. Adopting the evidently sensible epistemic principle to 'suspend belief until you get more information' therefore cannot provide a full answer to the question of what you should do when you don't know what to do.

To capture this, we might distinguish between moral propositions that agents *believe* they have from those that they *assume*. To assume a moral proposition, as we will understand this notion, is for it to play a role in your practical deliberations and to provide you with a motivating reason to act in certain ways.

Imagine that, after careful reflection, Jane thinks it 60% likely that abortion is permissible. Should she then, for the purposes of practical action, assume the proposition that she is permitted to have an abortion? Perhaps. Most applied ethics appears to regard this as the end of the matter. But some theorists disagree. For example, as Dan Moller argues, if Jane is only 60% convinced that abortion is morally permissible, then she has a 40% credence in the proposition that abortion is *not* morally permissible. Imagine that Jane thinks that *if* abortion is impermissible, it is akin to killing a toddler. That means she thinks that if she has an abortion, then by her own lights there is a 40% chance she is doing something very seriously morally wrong. Moral Uncertainty Theory can take this into account (Moller 2011).

We can make a useful, though perhaps imperfect, analogy with empirical uncertainty here. Imagine that Jane is uncertain about whether or not there is a child in her currently empty swimming pool. If she fills it up, then if there is a child in there, it will drown. Jane thinks there is a 40% chance there is a child in her swimming pool. Jane should, of course, take this into

account in deciding what to do: there is a 40% chance that she will do something seriously objectively morally wrong. But the same is true, according to Jane's own beliefs, in the abortion case.

When we assess theories of what we ought to do under moral uncertainty, it's important that we don't criticize them *simply* because they allow conduct that we think is morally forbidden (or forbid conduct that we think is allowed). *That* type of claim is the type to be considered in first-order moral deliberation, when we try to decide what is objectively right and wrong, permissible and impermissible. In judging whether a theory of moral decision-making under uncertainty is plausible, we must instead ask whether it is plausible that '*given this agent's beliefs*, it is reasonable for her to *assume x* is permissible (or impermissible) *when choosing what to do*.'

In this chapter, we seek to provide an overview of how different approaches to Moral Uncertainty Theory might inform various debates and issues within criminal law theory. Our discussion will be tentative for three reasons. First, the *point* of Moral Uncertainty Theory is that it cannot *on its own* provide moral guidance – it seeks to help agents decide what to do given their own moral beliefs. So, what Moral Uncertainty Theory 'has to say' about issues in the criminal law will depend upon what the 'inputs' are: what moral views people have, and what credence they have in them. Second, there remains debate and controversy about what is the right approach within Moral Uncertainty Theory. That is, there is disagreement about how we should take into account the various first-order moral beliefs that an agent has in deciding what they ought to assume for the purposes of action. The literature is in its infancy (the first systematic treatment is Lockhart 2000) and questions, issues, and problems are still emerging. Third, Moral Uncertainty Theory applies in the first instance to individual moral agents, providing guidance on what they should regard as morally best or morally permissible given their moral beliefs. But when we are considering issues of criminal law, we face social decisions about how to collectively design shared social institutions. That is, it is not immediately clear what the analogue of the individual agent is, or how to represent different credences in propositions concerning the moral status of different kinds of conduct when it comes to collective decisions. In many ways, our chapter is an exploration of the difficulties and complexities of trying to apply this approach to applied ethics to issues within the criminal law.

We shall proceed as follows. In the following section, we introduce Moral Uncertainty Theory in more detail, explaining some different approaches to the questions of what we ought to do when we don't know what to do. In Section III, we look at the criminal justice process,

and identify some key actors who might be morally uncertain, and the decisions about which they may face moral uncertainty. In Section IV, we focus in on questions about criminalization – how can morally uncertain individuals or collectives use Moral Uncertainty Theory when deciding how to vote on criminalization matters? In section V, we discuss the issue of how to compare the wrongs of impermissible criminalization and impermissible failures to criminalize. In Section VI we discuss sentencing. In Section VII we conclude.

II Moral Uncertainty Theory: Approaches and Problems

It is only recently that philosophers have begun to systematically turn their attention to what we ought to do under moral uncertainty.[3] Their discussions have, broadly-speaking, addressed three questions. The first concerns the validity and the nature of the ought in question. As we have phrased it, the question which Moral Uncertainty Theory addresses is *what should I do, given that I don't know what I should do?* Some have denied that this question makes sense. That is, they have denied that there can be an 'ought' which is sensitive to our moral beliefs – morally speaking, we simply ought to do what we ought to do! (Harman 2015; Weatherson 2002, 2014). Others believe that there *is* an ought here, but there is some disagreement about what *kind* of ought it is – a rational ought; a moral ought; a sui generis kind of ought? We'll largely set these controversies aside in this chapter.

The second question is substantive: what should I do when I don't know what I should do? Here, philosophers have sought to provide theories or frameworks into which agents can plug their first order moral beliefs, and emerge with an answer about what they ought to do, relative to their moral beliefs. Elsewhere (Barry and Tomlin 2016), we have pleaded for a little more precision about exactly what is being asked here. At the first-order moral level, we often make important distinctions between what an agent is permitted to do, what she is required to do, what would be morally best for her to do, and what she should do all-things-considered. All of these first-order moral questions have their moral-belief-relative analogues, and so the question 'what should do given that I don't know what I should do?' is underspecified. Our own interest, both elsewhere and here, concerns moral permissions, and so our question is: 'what am I (or we) permitted to do, given that I (or we) don't know what I (or we) are permitted to do?' We can call this the question of moral-belief-relative permissibility (MBR permissibility). Why focus on permissibility? Well, judgements of impermissibility carry a special kind of moral gravity, and impermissible actions attract a special kind of moral blame and censure. This focus seems especially important when considering the criminal law. One possible response to actions that are impermissible is to attach legal sanctions to them,

including criminal punishment. Imposing criminal sanctions is not a response that is ordinarily thought to be appropriate when it comes to conduct that is merely morally sub-optimal. And when it comes to punishing conduct, we want, first and foremost, to ensure that our punishments will be permissible: the philosophical debates about punishment turn on if, when, and why it is permissible, not whether it is admirable, or supererogatory. The issue of permissibility under uncertainty is, moreover, of special concern when thinking about what kinds of conduct to criminalize, how criminal procedure should be designed, which punishments are proportionate, and so on. We will canvass three general approaches that have been taken to MBR permissibility in the literature below.

The final question that philosophers have considered concerns inter-theoretic comparisons of value. The issue here is that most approaches to the question of what we are MBR permitted to do require us to take into account the recommendations or requirements of several moral theories. We can understand a theory as a set of principles from which we can derive different moral scores for conduct options. But it isn't clear that these theories are trading in a single 'currency', and thus whether they can be compared in the right way. We are going to set this issue aside in this chapter.[4] However, we think it's important to remember that while the comparisons that must be made are inter-theory, they are *intra-agent*. That is, the individual who is morally uncertain may be able to do the appropriate comparisons (presumably she knows how valuable or disvaluable the options are according to the moral reasons that generate her uncertainty), even if the moral theories *themselves* don't give her the resources for doing so.[5]

It must be acknowledged at the outset that any talk of the theories that an agent has credences in idealizes quite significantly ordinary moral deliberation in a few ways (Unruh n.d.). First, agents do not typically think of themselves as torn between two conflicting *theories* when making choices (they may not even have a clear idea of what a moral theory might look like). Second, agents are unlikely to articulate their commitments in terms of precise credences in rival moral theories.[6] Third, agents don't typically deliberate on the assumption that specific theories they have some credence in value different options in some very precise manner (assigning a 'moral score' to, say, to how the theory treats infringing someone's property rights to help someone else in need of shelter). Still, such idealisation can be illuminating. It *is* common for people to be in conflict about the moral status of different conduct options, and their reasons for this can be plausibly represented in terms of competing values to which they are drawn (or arguments they find convincing), which in turn can be seen as embodying competing theories. And, surely, we do have at least some intuitive idea, when we are feeling

conflicted about the moral status of some conduct option (whether it is permissible to eat meat, terminate a pregnancy, harshly criticize a colleague on social media), about our level of confidence in its permissibility. Finally, we also seem to have some intuitive idea of just how bad some conduct would be, according to some values we are committed to or some argument we are drawn to.

Three Approaches to MBR Permissibility

<u>My Favorite Theory</u>

How, then, can we tell whether some course of conduct will be MBR permitted? We will here consider three answers to this question. The simplest view, and the view which much applied ethics tacitly appears to endorse, is that you should follow a course of conduct which the theory in which you have the most confidence deems permissible. That is, once Jane has decided that she is 60% confident in a theory which says that she is permitted to have an abortion, then she should regard herself as permitted to have an abortion. Ted Lockhart (2000, 42-43) calls this the 'My Favorite Theory' approach to moral uncertainty. This approach provides consistent advice for agents: regard as permissible whatever, according to the theory you have most credence in, takes to be permissible. In addition, it does not require the agent to make intertheoretic comparisons of value, since she need only pay attention to how one theory—the one she has the most credence in— values the conduct options she is considering.[7]

The My Favorite Theory view has been subject to serious criticism (Lockhart 2000; Barry and Tomlin 2016; Bykvist 2017). One problem with My Favorite Theory is that it seems unduly sensitive to the individuation of moral theories. Say that in our abortion example, Jane is 60% confident in the theory that says she is permitted to have an abortion (Theory 1), and 40% confident in the theory that says she is not (Theory 2). According to My Favorite Theory, she is MBR permitted to have an abortion. Now imagine Jane discovers an important distinction within the way that the permissibility of abortion could be established. She is unsure whether, if it is permitted, it is permitted because the foetus lacks the requisite moral status, or because there are circumstances in which we are permitted to kill beings with moral status. She is equally confident in each of these views. So, Theory 1 is split into Theory 1a and Theory 1b. Jane now has 30% credence in each of these views. Jane's Favorite theory now turns out to be Theory 2, in which she has 40% credence. By splitting Theory 1 into two distinct theories, even though her credence in the permissibility of abortion remains constant, it becomes MBR impermissible for Jane to have an abortion. However, the basic intuition behind the My

Favorite Theory approach might be preserved by shifting from talk of theories to talk of options. That is, it might plausibly direct the agent to regard as MBR permissible any conduct that is such that she has more credence that it is permissible than that it is impermissible.[8] This would have the implication in the case above the Jane regards abortion as MBR permissible, even though she is unsure just which theory gives the best account of its permissibility.[9]

A much deeper worry about the My Favorite Theory approach is that it seems to dismiss the idea that moral risk ought to affect what we do, even when the stakes are very high (as they are in Jane's case in the example above.) Imagine an agent faces a moral choice between A and B. She has credence in two theories, Theory 1 and Theory 2, and has slightly greater credence in Theory 1. Theory 1 says A and B are both permitted. According to My Favorite Theory, the agent should then simply choose between A and B, and can do so on non-moral grounds – both are MBR permitted. Imagine the agent is personally indifferent between A and B, and so flips a coin. According to My Favorite Theory this is fine *even if* Theory 2, in which the agent has almost as much confidence as Theory 1, says that B is a very serious moral wrong akin to murder. Surely, when the agent is indifferent between the options, and one carries a high chance of being a serious objective moral wrong, then even if this option would confer a significant personal benefit to her, she has strong reason to choose the other option.

The above argument relies on the My Favorite Theory Approach appearing to deliver the wrong verdict about MBR permissibility. It is important to remember that when we consult intuitions about these cases, we should not be asking ourselves what *we think* is objectively permissible when it comes to the question of abortion, or even what *we* think is MBR permissible, but rather what *the agent who is uncertain* should regard as MBR permissible *given her beliefs*.

Expected Moral Value

The Expected Moral Value approach to Moral Uncertainty Theory addresses what we suggested was the principal defect of the My Favorite Theory approach. That is, the Expected Moral Value approach takes seriously not only the agent's credence in various theories, but also the moral value (and disvalue) that such theories attach to different conduct options that the agent is considering. The Expected Moral Value 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 (Lockhart 2000; Ross 2006). So to return to our running example, suppose that Jane has a 40% credence in Theory 2 according to which abortion is

akin to killing a toddler and a 60% credence in Theory 1 according to which abortion is permissible. While on the My Favorite Theory approach Jane should assume the view that she can permissibly terminate her pregnancy, the Expected Moral Value approach will imply that she should regard abortion to be impermissible: the expected moral disvalue of terminating her pregnancy is very high, since by her lights she would be taking a 40% of doing the moral equivalent of killing a toddler. We wouldn't ordinarily perform some action that carried a 40% chance (or, for that matter, a 20% chance) of being a serious moral wrong (unless moral reasons concerning the alternatives rendered that the best of a very bad set of options), even if that action would produce a significant personal benefit to us.[10]

One of the appealing features of this approach to moral uncertainty is that it treats the requirements of rational choice under moral uncertainty as broadly continuous with those that apply under empirical uncertainty. Still, the very features that make this seem a plausible approach in some cases lead to counterintuitive implications in others. The central problem with the Expected Moral Value approach is that it will always prioritize certainly permissible options over possibly permissible options. Imagine that, instead of a 40% credence in Theory 2 (according to which abortion is a very serious moral wrong), Jane has only a 0.1% credence in it. What should she regard as her permissible options in this case? Well, she knows that according to one of her theories (Theory 1—the one she has a 99.9% credence in), it is permissible either to terminate or to carry to term, while according to the other, more restrictive theory, only one of these options is permissible. So, given that each theory will regard carrying to term as permissible, while one regards terminating the pregnancy as a serious moral wrong, this approach will deliver the verdict that it is MBR impermissible for Jane to have an abortion. That is, this approach will often entail 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, even if we have *very* little credence in it, and even if (unlike in our case) the potential moral wrong is not a very serious one. We call this a 'nesting' case, since one theory (the more restrictive theory) 'nests' inside the more permissive theory, in that each and every conduct option permitted by the more restrictive theory is also permitted by the more permissive theory, while the permissive theory allows additional options.

Another challenge for the Expected Moral Value approach is posed by what we call 'Venn cases'. These are cases in which two (or more) theories agree on the permissibility of a limited set of conduct options (making them certainty permissible) but there are other conduct options which some of the rival theories regard as permissible but others do not (making them possibly permissible, from the agent's point of view). The Expected Moral Value approach

says only the certainly permissible options should be considered MBR permissible. The Expected Moral Value approach, however, pays no attention to *why* a given theory finds some conduct permissible, and ignores the possibility that moral theories may care about the *combination* of permissible options.

Consider the following example. Richard has thought long and hard about the permissibility of suicide. He is certain it is permissible, because he believes that people ought to, insofar as is possible, have control over when and how their life ends. Richard knows that next week he will face a moral dilemma. He is uncertain as to whether to visit a friend or a stranger in hospital, given that visiting the stranger will produce more overall wellbeing. Given that he is unsure, Richard will risk a minor moral wrong whoever he visits. However, since he is certain that suicide is permissible, the Expected Moral Value approach states that Richard is MBR *required* to kill himself – it is the only certainly permissible option. Much of the appeal of the Expected Moral Value approach strategy of always ranking certainly permissible options over possibly permissible options appears to rely on the idea that no theory 'loses out' when an agent's MBR permissible options are restricted to the options that all the theories endorse as permissible. We think, however, that this fails to take into account a potentially 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. Richard, for instance, has no credence in theories that say that suicide is morally required. He has credence in theories that value the combination of the options of committing suicide or not, and these theories *do* lose out if his MBR permissible options are restricted to the (objectively permissible) option of committing suicide.

Evaluating Option Sets

In sum, the central problem with the Expected Moral Value approach is that it looks at conduct options one by one, and pays no attention to the fact that more permissive moral theories don't only care about the permissibility of individual options, but the *range* of options open to the agent. For example, theories that support pro-choice care about the agent having the *choice* of whether or not to terminate their pregnancy, and Richard cares about having the *choice* of when to die. A MBR *requirement* not to abort, or to kill oneself in order to avoid risking minor moral wrongs, fails to take adequate account of these concerns.

An approach to Moral Uncertainty Theory that we have defended seeks to avoid these problems by switching the unit of concern from individual *conduct options* to *option sets* (Barry and Tomlin 2016). Option sets are combinations of options which, if accepted, are all

permissible. To take an example from first-order deliberation, imagine that Nina is deciding whether to go to the cinema or stay at home tonight. Both of these options are objectively permissible, and so her moral option set includes both going to the cinema *and* staying at home.

To explain how our approach is distinctive when applied to moral uncertainty, imagine that there are basically three conduct options available to an agent – A, B, and C. The Expected Moral Value approach asks us to look at A, at B, and at C independently, and then choose the option with the best score. Our approach looks at, and evaluates, the different potential *option sets* available to the agent. These include A, B, and C, but also AB, BC, AC, and ABC. What, we ask, should the agent regard as her MBR permitted *range* of options, and our answer to that question takes into account both the moral risks of individual conduct options (as the Expected Moral Value approach does) but also how the *range* of options looks. Essentially, we add an additional layer of scrutiny to the Expected Moral Value approach by morally evaluating combinations of options, and we think it will be sometimes worth accepting a small risk of objective impermissibility for a better option set.

To return to the example of Richard, Richard has three options: visit his friend, visit the stranger, or kill himself. These could be combined into different option sets: an option set that *only* includes killing himself, which, if accepted, would mean he is MBR required to do so; an option set that gives him a choice between all three options, which, if accepted, would mean he is MBR permitted to do any of them; and various combinations of two options. The option set that *only* includes suicide does very well in terms of individual options (no risk of objective wrongdoing) but *very* badly as an option set – it doesn't give him any other options, and he only values the option of suicide when and because it is available *alongside* the option *not* to commit suicide. We think Richard should be guided toward a different option set, even though the other option sets risk objectively impermissible conduct. He is MBR permitted to risk some objective wrongdoing.[11]

III. Uncertainty in the Criminal Law

Moral uncertainty can occur at many points during the criminal justice process. Consider this, somewhat simplified, model of how some particular person ends up being convicted of, and punished for, some particular crime.

I.   Voters elect representatives based, in part, on what they (or their parties) have conveyed about their beliefs concerning issues around the criminal law.

II.   Elected representatives choose to make Conduct C a crime – Crime C.

III.   Police officers enforce Crime C, arresting individuals they suspect of Conduct C. They arrest Bob.

IV.   Prosecutors decide to prosecute Bob for Crime C.

V.   A jury decides to convict Bob of Crime C.

VI.   A judge sentences Bob, handing down Punishment X for Crime C.

VII.   Prison wardens and officers must carry out Punishment X.

VIII.   A Parole Board decides to release Bob early, before all of Punishment X has been completed.

The actors involved in all of these decisions may experience moral uncertainty (as well as empirical uncertainty, which, for simplicity we will set aside here). Citizens can be unsure about what general approach to the criminal law is the correct one. Legislators can be unsure about whether to criminalize Conduct C. Police – as individuals and as institutions – are given discretion about what laws they prioritize the enforcement of, and, relatedly, how to allocate spending, and may be unsure about which offences to enforce, or which to prioritize. Prosecutors are given discretion over whether to press charges, in particular concerning whether or not prosecuting the alleged offence would be in the 'public interest'. They may be morally uncertain about what is in the public interest. Judges may be morally uncertain as to which of the various punishments they could hand down is morally appropriate. Prison wardens may be morally uncertain about whether they are being too harsh, or too lax, on their prisoners. The Parole Board may be morally uncertain about whether or not an offender's behaviour warrants early release.

In addition, at each of these points, we can distinguish two main ways in which moral uncertainty theory might be relevant to the criminal law. First, we can consider moral questions that arise *within* any standing practice of criminal law. That is, judges, lawyers, jurors and others involved in the criminal process face practical decisions about what to do, where they can be morally uncertain about what to do. For example, a prosecutor must decide whether to bring a case; or a judge must decide what sentence to pass. Second, we can consider moral questions concerning the design of the criminal justice system itself. For example, we must individually and collectively make decisions about what kinds of conduct should be criminalized, which punishments can attach to which criminal acts, how criminal process should be structured, what evidential standards to employ, and so on.[12] These two sets of questions are of course related. For example, a juror in a criminal trial may be uncertain about

the moral character of the criminal legal process, or about how to act when you believe certain aspects of such processes are unjust.

IV. Uncertainty and Criminalization

Let's turn to how Moral Uncertainty Theory might be brought to bear on questions concerning the design of criminal justice institutions. To sharpen the focus of our discussion, we'll consider the question of what kinds of conduct to criminalize. Note first that there are some important differences between the question that faces us here, and the question that is tackled, in general, in the Moral Uncertainty Theory literature. The question that Moral Uncertainty Theory is typically concerned with is '*given that I am uncertain, what am I permitted to do?*' When we're thinking about criminalizing some conduct, however, the question is '*given that we are uncertain, what are we permitted to criminalize?*' This involves two important additional complexities that the first question does not involve.

The first complexity arises from the move from *I* to *we*. Sometimes we might be asking a question about criminalization under moral uncertainty as an individual – either as an individual citizen or legislator – but sometimes we might be asking it as a group – as a legislature, or as a society. What would it mean to say that we as a group regard some question regarding criminalisation as uncertain? It could refer to the fact that each and every individual comprising the group is individually morally uncertain about the conduct. Or it could mean that, while some individuals may indeed be morally certain about the issue, there is no strong consensus on it. That is, the moral status of the conduct is not treated as morally *settled* by the group but is instead regarded as something about which there is reasonable disagreement.[13] This poses important, and complex, questions about what theories we should be taking into account, and what credence we should assign to them. When, as a society, we debate criminalizing some conduct, should I, as a member of society, be trying to figure out what *we* are permitted to do, given that *we* are uncertain, thereby taking into account the diversity of beliefs concerning the moral status of that conduct? Or should I be trying to figure out what legislation to fight for, given that *I* am uncertain about the moral status of some conduct, and so taking into account *my* beliefs?

A second issue prompted by the move from *I* to *we* relates specifically to our own preferred approach to Moral Uncertainty Theory (Evaluating Option Sets). Recall that our approach is concerned that agents should have a *range* of conduct options. But people who have these kinds of concerns about *individuals* often do not have them, or at least have them in the same way, about political institutions. We might be concerned if morality leaves individuals with no space to shape their own lives and make their own choices, and it is this kind of concern that can only

be incorporated if we focus on option sets instead of individual conduct options. But when we ask about whether or not to criminalize some conduct, we ask about what *we* should do *as a political community*. And as Thomas Nagel memorably put it, 'institutions, unlike individuals, don't have their own lives to lead' (Nagel 1991, p.59). So it is unclear whether our particular approach to moral uncertainty theory will have much purchase at the institutional level. However, it is possible that we ought to allow political communities leeway to shape themselves through their criminal law. MBR permissibility, of course, already does this – the society's moral values will shape their criminal law. The question is whether, beyond this, they should have scope to choose their criminal laws, beyond the demands of moral-belief-relative morality. Of course, decisions about whether to criminalize conduct need to be sensitive to beliefs about the importance of individuals having a range of options. After all, criminalization limits the freedom of individuals by attaching legal and social consequences to some options. But should we also be concerned that the *group* retains freedom to devise its own criminal laws, and that morality must not demand too much of *it*?

The second complexity introduced by considering criminalization is that decisions of criminalization involve, essentially, evaluating two kinds of conduct. That is, we need to ask ourselves both whether the conduct to be criminalized is morally impermissible *and* whether our criminalizing it is morally permissible. We can be morally uncertain about both. We can be uncertain both about whether some conduct is impermissible, and about whether moral impermissibility is a necessary condition of criminalization. We can also be uncertain about what kind of wrongfulness that matters when deciding whether to criminalize conduct— if the conduct is MBR impermissible, is that enough to potentially warrant criminalization, or is it objective moral impermissibility that matters here? If some conduct is criminalized on the basis of its MBR impermissibility, is that ok, or should we factor in the chance that it *might* be objectively permissible?

To illustrate these difficulties, consider again the issue of abortion. Suppose an agent has a 40% credence in a Theory 1 according to which abortion is impermissible, and a 60% credence in a Theory 2 according to which it is permissible. Nevertheless, the agent is certain that we should never criminalize conduct that is morally permissible, but holds that it is sometimes permissible to criminalize conduct that is morally impermissible. How then should an agent with such credences think that *law* should regard this conduct? Note that this a somewhat different question than how the agent should regard the question of the moral permissibility of abortion, given her moral uncertainty. Still, if we hold the My Favorite Theory approach, the answer is probably straightforward: she shouldn't regard it as permissible to criminalize abortion since, by

the lights of the theory in which she has the highest credence, it is a morally permissible practice (so it is a MBR permissible practice, and therefore she should regard it as MBR morally impermissible to criminalize the conduct given her belief that it is always impermissible to criminalize permissible conduct.)

While My Favorite Theory is a controversial view about moral uncertainty, it does seem to capture roughly how representative democracies *do* treat the question of criminalization in practice: the majority's favorite theory (which might not be the theory with the highest average credence) wins the day. Unless some measure to criminalize conduct runs up against a constitutional provision, a bare majority is often enough to pass legislation that criminalizes some conduct. Moreover, there are no formal or informal norms about what kinds, or strengths, of beliefs individual legislators ought to have before voting to criminalize some conduct. So as long as more than half of them are convinced enough to decide to vote to criminalize the conduct, it will be criminalized, regardless of how uncertain various legislators or voters are, or what the moral risks are if they incorrect about this. For example, if individual legislators vote to criminalize some conduct on a balance-of-probabilities basis, and bills pass on a bare majority basis, then the conduct will be criminalized if 51% of legislators have a 51% credence in the proposition that the conduct can permissibly be criminalized *even if* the other 49% of legislators have no credence in this proposition at all. Therefore, conduct can be criminalized even when the average credence in the permissibility of doing so is 26%. In our view, this seems a very good reason to depart from common practice and reject the My Favorite Theory approach to criminalization (see further Tomlin 2013).

If we apply the Expected Moral Value approach to the question of criminalization things are more complex. There are broadly two different approaches we could take to the criminalization of conduct under the Expected Moral Value approach. On the first, which we can call the *two-step approach*, we *first* ask 'is the conduct MBR permissible?' and then, second, we address the question of criminalization. In the case of abortion, then, this two-step approach would first have us ask 'according to Expected Moral Value theory is abortion MBR permissible or MBR impermissible?' Let's imagine that, according to Expected Moral Value theory (with a given set of inputs), abortion is MBR impermissible. The two-step approach would then have us ask 'given that abortion is MBR impermissible, is criminalizing abortion MBR permissible or impermissible?' If the only relevant question was the MBR moral status of abortion (which it surely isn't), then it would be MBR permissible to criminalize abortion.

An alternative approach is the *one-step approach*. According to this approach, we should ask *directly* 'is criminalizing abortion MBR permissible?' and then work out the relevant moral

scores of criminalizing the conduct and not criminalizing the conduct, including the probabilities of abortion turning out to be objectively permissible and impermissible.

Essentially, this would involve constructing a two-by-two matrix. On one axis we would have the decision to be made – criminalize, or don't criminalize. On the other, we would have our moral uncertainty – is abortion permissible or impermissible?

One risk is that Theory 1 is correct. If we make abortion legally impermissible, then this would involve criminalising a practice that is in fact morally permissible. This would be in contravention of our stated principle (which, in this example, the agent is fully confident of), that we may only criminalize conduct that is morally impermissible. We would then need to assess how *serious* a wrong the act of criminalizing morally permissible conduct would be. To some extent, this will depend on the magnitude of the criminal sanction. Generally, however, criminalizing morally permissible conduct would be very seriously wrong indeed, given some of the distinctive features of the criminal law. In criminal trials, defendants are faced with the prospect of hard treatment: losing their rights, liberties and, in some jurisdictions, even their lives. Moreover, by identifying them as 'criminal offenders', society expresses an attitude towards them – that their conduct is paradigmatically blameworthy – that can be seriously damaging to them. Criminal sanctions thus stigmatize the offender and substantially affect the way they are treated within their communities. If the criminal law possesses this strong expressive function, convicting an agent of a criminal offence when they engage in morally permissible conduct may carry a morally relevant cost well above and beyond the cost to the person convicted. This is because the agencies that purport to be acting in the name of justice are imposing unjust harm. Unjust harm imposed on people through the agency of the state is often taken to be of particular moral consequence. Many people care that the state agencies that represent them not only bring about desirable end-states, but that they take special precautions not to impose unjust harm on individuals through what they do, even when this may lead them to allow still greater harms brought about by others (Barry 2005; Enoch 2007; Hosein 2014). That is, many people are not morally indifferent to the distinctions between what a state brings about directly through its agencies and what it fails to prevent. The idea that someone would be criminally sanctioned in the name of justice and under the color of law for acts that are morally permissible is deeply disconcerting (Husak 2007; Tomlin 2013). And of course there are other significant costs to the population when conduct is wrongly criminalized. Most importantly, it is not only those who nevertheless commit the act and are punished who are wronged. At the very least, those who *would have* committed the act but refrain from doing so because of the threat of criminal sanction can also be wronged. Plausibly, all those who *could have* committed the act

are wronged – even if they have no desire to perform the act in question, since they are threatened with sanction if they do so, even though they ought to be able to perform the act in question. How serious this wrong would be would vary from case to case, but in a case featuring our own bodies, and our core rights, it would be a very serious wrong indeed.

These considerations point towards an important disanalogy between how an individual agent reasons about the MBR permissibility of *her* terminating a pregnancy and how *we* as members of a society reason about whether to criminalize abortion. The individual agent must consider how to weigh the moral risk of terminating her pregnancy (insofar as she has some credence in a theory according to which abortion is a serious moral wrong) against the costs that *she* and others will face if she carries the foetus to term. However, when we deliberate about whether to criminalize abortion, we are considering how to weigh the moral risk of making a morally impermissible practice lawful against the costs that *we* will impermissibly impose on others (namely, women who might consider terminating their pregnancies but will be prevented from doing so.) Therefore, criminalizing abortion carries huge moral risks, and agents should be very wary of proposing it, *even if, when they individually deliberate, they come to regard abortion as MBR impermissible*.

If Theory 2 is correct, and abortion is objectively impermissible, then one moral risk would appear to be in *not criminalizing* the conduct. It is, however, questionable whether *not criminalizing* objectively impermissible conduct is impermissible, and therefore whether it poses a moral risk at all. Criminalization theory is largely concerned with providing an account of the conditions under which the criminalization of some conduct is *permissible.* It often proceeds by way of setting up hurdles to criminalization, and criminalization is regarded as permissible if and only if it manages to clear all of the hurdles (see, for example, Husak 2008). For example, there is extensive debate about how best to characterize the harm principle—roughly, that the purpose of the criminal law is to prevent harm— but most versions of the harm principle see it as a *necessary condition* on criminalization (Tomlin 2014a, Tadros 2011, Edwards 2014). This is of course understandable – given our ever-proliferating criminal law, politicians and states hardly need to be given reasons to criminalize (Husak 2008, ch. 1), and so the focus has been on when criminalization *shouldn't* occur. But when we come to Moral Uncertainty Theory, the difference between it being permissible to criminalize some conduct and it being not only permissible but also required becomes very important. That is, if criminalization is merely *permissible*, then our running example is a 'nesting case' – if Theory 1 is correct (e.g. abortion is objectively permissible), criminalization will be a serious moral wrong; if Theory 2 is correct (e.g. abortion is

objectively impermissible), then criminalization and not criminalizing may both be permissible. Recall that the Expected Moral Value approach is excessively conservative when it comes to risking moral wrongdoing – certainly permissible conduct is *always* favoured to possibly impermissible conduct. This is why the Expected Moral Value approach seems destined to say that, where there is any doubt about its permissibility, abortion is nearly always MBR impermissible (since carrying to term is almost always permissible). But for the same reason, if criminalizing some conduct is, at best, merely permissible according to both theories we have credence in, but also carries a risk (however small) of being impermissible, while not criminalizing is certainly permissible, then the Expected Moral Value approach would say we are MBR required to not criminalize the conduct.

Because of this kind of case, Moral Uncertainty Theory might play a substantial role in shaping the direction of first order moral theorizing about the criminal law. As observed above, moral theorizing about the criminal law often takes the form of proposing, defending, and examining justificatory hurdles to the *permissibility* of criminalization (for an extremely prominent example, consider the collective title of Joel Feinberg's four-part classic – *The Moral Limits of the Criminal Law* (Feinberg 1984, 1985, 1986, 1988)). But if criminalization is only ever, at best, permissible, then the Expected Moral Value approach to criminalization under moral uncertainty will only allow criminalization when it is *certainly* permissible. Things would be different, however, if criminalization were, at least sometimes, morally *required*. In that case, we would act impermissibly in *failing to criminalize* some conduct. This seems plausible with respect to certain kinds of conduct. For example, a state that failed to criminalize the killing of innocent, non-threatening people would seem to wrong its citizens quite seriously – but we need criminal law theory to tell us more about when and why we are required to criminalize.

If, under Theory 2 (according to which abortion is impermissible), we would act wrongly by *failing* to criminalize abortion, and under Theory 1 (according to which abortion is permissible) we would act wrongly by criminalizing abortion, then we do not have a nesting case. Instead we have a 'cross-condemning case,' in the sense that we risk acting wrongly whatever we do. Given this, we must assess the moral disvalue of both options, weight them by their subjective probabilities, and see which has the best expected moral score.

We have already looked at the wrongs of criminalizing abortion should it turn out to be objectively morally permissible. What might be the wrongs of *not* criminalizing abortion if it were objectively morally impermissible? Now, one might think that, according to Theory 2, the costs of legally permitting conduct that it regards as impermissible would be negligible. After all, we don't think that all behaviour that is morally impermissible (or even seriously wrong)

should be subject to criminal sanction—the betrayal of spouses or the failure to deliver on crucial promises, even some that we are contractually obliged deliver on, are cases in point. And sometimes this seems plausible for all of the reasons mentioned above—the moral risks of overcriminalization are typically quite high, whereas the risks of not criminalizing conduct that is morally impermissible may not be so great. However, this will not always be the case. More specifically, if Theory 2 holds that abortion is akin to murder then it would also regard practices that permit abortion to be at least morally comparable to practices which permit murder. Indeed, advocates of Theory 2 may plausibly claim that the criminal law is not only expressive in what it forbids but in what it *permits*. And if this is so, then it would attach a great deal of disvalue to the option of failing to treat abortion as a criminal offense, unless doing so carries very high counter-veiling costs.

Still, there might still be important disanalogies between murder and the termination of pregnancy that could change the calculus even if Theory 2 is correct concerning the moral status of abortion. For one thing, there is a great deal of disagreement about abortion, and this might prove relevant to the criminalization question, even at the level of objective morality. For example, that killing innocent, non-threating adults is forbidden is clearly regarded as settled by the society in the way that the practice of abortion is not. This could make a difference to how adherents of Theory 2 might judge the wrongfulness of the conduct of those who terminate their pregnancies, since those who do terminate their pregnancies do not in all likelihood consider this practice wrongful, much less akin to murder. This is where the issue of collective uncertainty may become relevant. A theory could take the cost of imposing criminal sanctions on conduct that, while impermissible, is the subject of a great deal of moral disagreement to be quite substantial. (This might explain why some ethical vegetarians, for example, would regard it as impermissible to impose criminal sanctions on meat eaters even though they regard eating meat as a serious moral wrong.)

Our own approach to moral uncertainty, Evaluating Option Sets, would differ from the above analysis only in that it would be less likely to regard abortion as MBR impermissible in the first place, since, as we observed above, our approach does not always forbid possibly permissible options when certainly permissible options are on the table. This is because, at the individual level, it is concerned to give full voice to theories that care about choice. However, at the state level, the reasoning would likely be very similar to that under the Expected Moral Value approach, since these kinds of concerns about moral restrictiveness do not necessarily have purchase for institutions.

V. Comparing Wrongs

If an agent is uncertain about the moral status of abortion, and, if she thinks that if abortion were *definitely* impermissible the state would be *required* to criminalize it, then in deciding whether she should regard the criminalization of abortion as MBR permissible she must balance the potential wrongs of criminalizing abortion when it should not be, and not criminalizing abortion when it should be, in light of the theories in which she has credence. We have tried, above, to identify some of the considerations that these theories might regard as relevant. But how can we gauge their respective weights?

One possible clue is to be found within the prevailing norms of the criminal justice system itself (Tomlin 2013). Consider the 'beyond reasonable doubt' principle, which is the so-called 'Golden Thread' of English law. In Viscount Sankey's famous words: 'Throughout the web of the English Criminal Law one golden thread is always to be seen, that it is the duty of the prosecution to prove the prisoner's guilt ... no attempt to whittle it down can be entertained' (quoted in Ashworth 2006, p.83). Insofar as both of the theories in which the agent has some credence are committed to this principle, it may provide some guidance as to how agents with credence in these theories should weight the respective wrongs of incorrect criminalization and incorrect failures to criminalize. Of course, the beyond reasonable doubt principle concerns how *juries* should regard the *empirical questions* that relate to whether or not a defendant acted in a way that was legally prohibited. However, the grounding of this principle may be informative in terms of how to weight the moral risks of criminalizing under moral uncertainty (see Yaffe, this volume, for further discussion of the beyond reasonable doubt principle).

The basic thought here is that the beyond reasonable doubt principle suggests that when it comes to convicting accused people, we should be very careful not to convict the innocent, even if this means letting the guilty go free. Consider William Blackstone's famous claim, that it is better that ten guilty people go free than that one innocent suffer conviction and punishment (Blackstone 1765). Why is this so? It is surely because these people have done nothing to warrant punishment, and, for all the reasons enumerated above, we must be very careful not to inflict punishment on them.

If the reasoning underlying the beyond reasonable doubt is principle is correct, then any theory that is committed to it will treat punishment as a tool that should be wielded very carefully indeed. And this further seems to suggest that it will wield *criminalization* very carefully. After all, being punished for something that is not punishable is not so very different from being punished for something you didn't do.

While the reasoning underlying the beyond reasonable doubt principle can *inform* the way in which we balance the potential wrongs involved in criminalization, there are important differences between decisions not to criminalize conduct and decisions not to punish some individual for an existing crime. One difference relates to the expressive or declarative functions of the criminal law. Even if nobody is ever convicted and punished for Crime X, there might be an expressive or declarative value to having it on the statute books. Another difference concerns fair warning and reasonable avoidability. That is, those who are convicted of crimes they didn't commit had no reasonable opportunity to avoid conviction. Those who are convicted of crimes which should not be crimes potentially did have such an opportunity (though how 'reasonable' the opportunity is will depend upon the legislation in question and the background circumstances of the agent).

VI. Setting Punishments under Uncertainty

Before concluding, let's briefly consider some issues concerning the determination of punishments under moral uncertainty. Imagine that we are certain that Conduct C should be a crime, it is a crime, and Bob has been convicted of it. Bob now stands before the judge, who must decide on Bob's sentence. The sentence will be a combination of two sets of decisions – those made by politicians and other political actors who have established the range of sentences that the judge can lawfully impose for this crime, and the decision made by the judge in the individual case. These actors can be reasonably uncertain both about the general principles that ought to govern sentencing, and about how those principles apply to a particular kind of conduct, or the individual case.[14] Here, we must balance the risks of *overpunishing* against the risks of *underpunishing*. One possible analysis of this balance again draws on the beyond reasonable doubt principle: that principle appears to emerge from balancing failure to deliver warranted punishment against delivering unwarranted punishment. Overpunishment is a form of unwarranted punishment, underpunishment a failure to deliver warranted punishment. If the reasoning underlying the beyond reasonable doubt principle can inform sentencing decisions, it suggests that punishments at the lower end of the spectrum being considered should be preferred (Tomlin 2014b).

However, things are potentially trickier when the criminal prohibition in question is one that passed despite deep reservations of those supporting it. For example, imagine a society has a long debate about whether to introduce a 'good Samaritan' law, one which criminalizes the conduct of failing to assist others in cases of easy, non-costly rescue. In the end, the legislature decides that it will establish the good Samaritan law. They then must decide what

sentences can be handed out to those that breach this law. This raises two interesting issues – one of first-order morality, the other of Moral Uncertainty Theory. First, should the society's uncertainty and disagreement about the status of some law be taken into account in establishing the morally appropriate sentence for breaches of that law? The answer seems to depend on big questions concerning the point of criminal law and punishment, and we don't have space to do more than raise them here. For example, if we are straight-up retributivists, then the correct punishment will reflect the moral gravity of the act – that society is unsure about whether or not to criminalize the conduct is neither here nor there. Either it warrants criminalization and punishment, or it doesn't. On the other hand, more expressive or communitarian-oriented theories might see things differently: if sentencing is a form of *communal* sanction, it seems plausible that the sentence attached to the crime should reflect the unease with which the society introduced the legislation.

At the level of Moral Uncertainty Theory, we can also ask whether the uncertainty with which the society introduced the legislation should affect sentencing calculations. Essentially, when an agent is uncertain about attaching a sentence, we need to take account of all the different sentences that different moral theories would endorse. Let's say that when the good Samaritan law is introduced, a member of the Sentencing Commission is torn as to whether failures to assist in cases of easy, non-costly rescue are significantly less seriously wrong than murder (say, on the general ground that doing harm is much worse than allowing harm to occur, all else being held equal), or are morally equivalent to unlawfully killing innocent, non-threatening people. Therefore, she is unsure as to whether a sentence of 1 or 8 years is appropriate. In deciding what sentence to recommend, she must balance the moral risks of over- and under-punishment. But let's say that she, like the wider society, was also deeply conflicted about whether to introduce the legislation at all. That means that she also thinks that it is possible that such failures to rescue should not be criminally sanctioned at all. Should this be taken into account? If so, a one year sentence *also* risks over-punishment. If not, a one year sentence carries no risk of wrongdoing at all. Should sentencing under moral uncertainty proceed as if criminalization is warranted – it only remains to be seen what sentence to attach, or should it reflect the uncertainty with which legislation was introduced?

VII. Concluding Remarks

This chapter has explored how we should approach moral questions concerning the criminal law – in particular what conduct we can or should criminalize, and what sentences we can or should attach to criminal conduct – under moral uncertainty. Criminal law theory is hard, and

so we are often, and often *should be*, reasonably uncertain about what the morally right answers are concerning how the criminal law should be structured. And of course the stakes attaching to how we resolve these questions are very high – the criminal law, and the conduct it often responds to, can wreak havoc in people's lives, innocent and guilty alike. Using the criminal law inappropriately can have disastrous consequences. Failing to use it when we could do so fairly and effectively can also be a grave moral wrong.

Moral Uncertainty Theory is a growing area of applied ethics that seeks to help us to decide what to do when we are morally uncertain. However, we have seen here is that it is not easy to apply it to the criminal law. Moral Uncertainty Theory ordinarily deals with questions concerning individual morality. We've seen that when we try to apply it to questions of how collectively we should respond to breaches of inter-personal morality, things rapidly become complex. That is why this chapter perhaps raises more questions than it answers.

Nevertheless, we think Moral Uncertainty Theory holds promise for informing our thinking some of the most pressing questions in criminal law theory. For example, is it good enough that we always act on the theory in which we have the most confidence? Shouldn't we take account of the potential wrongs we will do if we criminalize conduct that ought to be legal? If we should, we have other questions to answer: What would those wrongs be, and how serious would they be? Is the state only ever permitted to criminalize conduct, or is it, at least sometimes, required? If it is required, when it is so required? What are the wrongs of the state failing to criminalize conduct it should have criminalized? How significant are they compared with those of incorrect criminalization? And, once conduct is criminalized, how should we approach sentence-setting under uncertainty?

Bibliography

Ashworth, Andrew. *Principles of Criminal Law*, Fifth Edition. Oxford: Oxford University Press, 2006.

Barry, Christian. "Applying the Contribution Principle," *Metaphilosophy*, 36 (2005): 210–27/

Barry, Christian, and Tomlin, Patrick. "Moral Uncertainty and Permissibility: Evaluating Option Sets," *Canadian Journal of Philosophy* 46 (2016): 898-923.

Blackstone, William. *Commentaries on the Laws of England*. Oxford: Clarendon Press, 1765–69.

Bykvist, Krister. "Moral Uncertainty," *Philosophy Compass* 12 (2017).

Edwards, James. "Harm Principles," *Legal Theory* 20 (2014): 253-285.

Feinberg, Joel. *The Moral Limits of the Criminal Law. Vol. 1, Harm to Others*. New York: Oxford University Press, 1984.

Feinberg, Joel. *The Moral Limits of the Criminal Law. Vol. 2, Offense to Others*. New York: Oxford University Press, 1985.

Feinberg, Joel. *The Moral Limits of the Criminal Law. Vol. 3, Harm to Self*. New York: Oxford University Press, 1986.

Feinberg, Joel. *The Moral Limits of the Criminal Law. Vol. 4, Harmless Wrongdoing*. New York: Oxford University Press, 1988.

Feinberg, Joel. "The Expressive Function of Punishment." In
Doing and Deserving: Essays in the Theory of Responsibility, edited by Joel Feinberg.
Princeton: Princeton University Press, 1970: 95–118

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

Gustafsson, Johan E., and Torpman, Olle. "In Defence of My Favorite Theory." *Pacific Philosophical Quarterly* 95 (2014): 159–174.

Harman, Elizabeth. "The Irrelevance of Moral Uncertainty." In *Oxford Studies in Metaethics*, Vol. X., edited by R. Shafer-Landau. Oxford: Oxford University Press, 2015.

Husak, Douglas. *Overcriminalization.* Oxford: Oxford University Press, 2008.

Husak, Douglas. "Ignorance of Law: How to Conceptualize and Maybe Resolve the Issue."
This Volume.

Jackson, Frank, and Smith, Michael. "Absolutist Moral Theories and Uncertainty," *Journal of Philosophy* 103 (2006): 267–283.

Lockhart, Ted. "A Decision-Theoretic Reconstruction of Roe V. Wade," *Public Affairs Quarterly* 5 (1991): 243-258.

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

MacAskill, William. "Moral Uncertainty as a Voting Problem," *Mind* 125 (2016): 967-1004.

MacAskill, William, and Ord, Toby. "Why Maximize Expected Choice-Worthiness," *Noûs* (forthcoming).

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

Nagel, Thomas. *Equality and Partiality*. New York: Oxford University Press, 1991.

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

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

Sepielli, Andrew. "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, 2009.

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

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

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

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

Tadros, Victor. "Harm. Sovereignty, and Prohibition," *Legal Theory* 17 (2011): 35-65.

Tomlin, Patrick. "Extending the Golden Thread? Criminalisation and the Presumption of Innocence," *Journal of Political Philosophy* 21 (2013): 44–66.

Tomlin, Patrick. "Retributivists! The Harm Principle is not for you!" *Ethics* 124 (2014a): 272-298.

Tomlin, Patrick. "Could the Presumption of Innocence Protect the Guilty?", *Criminal Law and Philosophy* 8 (2014b): 431-447.

Unruh, Charlotte. "Moral Uncertainty as a Real-World Problem," Unpublished m.s.

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

Weatherson, Brian. "Running risks morally," *Philosophical Studies* 167 (2014).: 141–163.

Yaffe, Gideon. "When Does Evidence Support Beyond Reasonable Doubt?" This Volume.

---

[1] 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.

[2] Abortion is an issue that many writers on moral uncertainty have focused on (see, for example, Lockhart, 1991, 2000, Guerrero 2007, Moller 2010, Weatherson 2014, Barry and Tomlin 2016), and of course its moral and legal status remain the subject of heated disagreement, so we will use it as our running example throughout this chapter.

[3] For a very useful overview of the debates, see Bykvist 2017. Some distinctive contributions to the literature include: Ross 2006; Moller, 2011; Jackson and Smith 2006; Guerrero 2007; Lockhart 2000; Sepielli 2006, 2009; Weatherson 2002; MacAskill 2016; Barry and Tomlin 2016; Gustafsson and Torpman 2014; MacAskill and Ord, forthcoming.

[4] For discussion see: Bykvist 2017, Lockhart 2000, ch.2, Gustafsson and Torpman 2014, Hedden, 2016, Ross, 2006; Sepielli, 2009.

[5] We are grateful for useful correspondence with Peter Vallentyne on this issue. This is not to say, of course, that the typical moral agent will actually engage in such deliberations.

[6] Indeed, any attempt to do so might be implausible. It seems plausible that our credences in theories, as in many other propositions, are rather imprecise or indeterminate.

[7] Gustafsson and Torpman regard this as a principal advantage of the MFT approach. Gustafsson and Torpman 2014, p.60.

[8] MacAskill and Ord refer to this as the 'My Favorite Option' approach. MacAskill and Ord, forthcoming, pp.9-10.

[9] For a more extended response to these worries about individuation see Gutafsson and Torpman 2014, pp. 171-2.

[10] This approach assumes, which the My Favorite Theory approach does not, that we can indeed make meaningful comparisons across theories in terms of how much they value various options.

[11] For criticism of our approach, see Ord and MacAskill, forthcoming. The main issue they identify is 'double counting' of demandingness or restrictiveness concerns. Our reply to this is outlined at Barry and Tomlin 2016, pp. 916-7.

[12] Some of the decisions we must make concern epistemic issues. We must, for example, consider whether ignorance of the law or uncertainty about what it requires can serve as an excuse for criminal acts. See Husak, this volume, for discussion of this issue. This type of uncertainty, about the law's requirements, is distinct from the uncertainty that we are focusing on here, which concerns what laws we *ought* to make.

[13] That the morality of some conduct is considered to be settled by a group is consistent with the fact that some people choose nonetheless to engage in that conduct—the killing of innocent, non-threatening people is an example.

[14] Our question here concerns how the moral uncertainty of those deciding the sentence should influence their decision. A further, potentially relevant, source of uncertainty is from the convicted party themselves – should the offender's legal or moral ignorance or uncertainty affect sentencing? On this issue, see Husak (this volume). Of course, those deciding the sentence may be uncertain over how to handle the offender's uncertainty, and this may influence their decision