In dubious battle: uncertainty and the ethics of killing

Seth Lazar[1]

Abstract How should deontologists concerned with the ethics of killing apply their moral theory when we don't know all the facts relevant to the permissibility of our action? Though the stakes couldn't be higher, and uncertainty is endemic where killing is concerned, few deontologists have an answer to this question. In this paper I canvass two possibilities: that we should apply a threshold standard, equivalent to the 'beyond a reasonable doubt' standard applied for criminal punishment; and that we should fit our deontological ethical theory into the apparatus of decision theory. I show that the first approach faces insurmountable obstacles, while the second holds much more promise for deontologists than they (and their critics) might first have assumed.

1 Introduction

Uncertainty is endemic to all action, but especially prevalent and troubling in the use of lethal force. The stakes are high. Time to think is short. The facts are disputed, inaccessible, often deliberately obscured. What, if anything, can moral philosophers tell those deciding whether to kill?

A terminological point: throughout this paper I use 'uncertainty' in its ordinary meaning to cover all species of imperfect information. I do not distinguish as economists and decision theorists often do between risk (when subjective probabilities are defined) and uncertainty (when they are not).

✉ Seth Lazar
sethlazar@gmail.com

1   School of Philosophy, Australian National University, Office 1107 Coombs Building 9, Canberra, ACT 0200, Australia

＠ Springer

We could tell them to *do what is objectively right*. And we can say what that would be, if the facts were thus and so. But since the facts are uncertain, we must say more. I think we should offer them, at least, a standard of subjective permissibility. And since the best accounts of the ethics of killing are deontological, we need one that works for deontological ethics.

Deontologists are sometimes accused of ignoring uncertainty entirely (Fried 2012). This is unfair (see Sect. 3 below). But deontologists have until now failed to defend a standard of subjective permissibility.[1] Sometimes consequentialists furnish one for them—but usually hoping either to refute deontology or to show that it is no more than a species of consequentialism (Oddie and Milne 1991).[2] In this paper, I contribute to the deontological response, assessing two deontological approaches to the subjective permissibility of intentional killing.

I focus on killing for three reasons. First, though few of us personally face life-and-death decisions, our governments take them every day and, in doing so, risk killing innocent people. The stakes could hardly be higher. This paper is anchored in concern for real practical problems, not only theoretical enquiry. Second, I think we cannot assess a standard of subjective permissibility apart from the theory of objective permissibility that it seeks to apply. So it helps to choose a problem for which we already have a persuasive theory of objective permissibility. And third, the best theory of the ethics of killing is deontological. If we cannot devise a standard of subjective permissibility that works for deontological ethics 'on its home turf', then there is little hope of doing so elsewhere.

The first attempt is the 'Threshold View'. It is not a complete standard of subjective permissibility. Instead it offers a necessary condition for an act of intentional killing to be subjectively permissible. However, since it reflects most deontologists' first thoughts about uncertainty, and since it is all deontologists have offered on this question thus far, it is important to see whether it succeeds. In Sect. 3 I argue that it does not.

The second attempt is more ambitious, though it still falls short of offering a standard of subjective permissibility. Instead, it offers a sufficient condition for an act's being subjectively permissible. Its structure is decision-theoretic. It tells deontologists to represent their moral theories in a way that assigns utilities to outcomes, and then (with some qualifications) maximise expected moral utility. This approach is typically floated by consequentialists (Broome 1991; Oddie and Milne 1991; Colyvan et al. 2010), so deontologists are understandably sceptical. I argue that their scepticism is unwarranted. One can adopt the decision-theoretic framework as a sufficient condition for subjective permissibility without abandoning any core deontological precepts. What's more, doing so allows us to avoid the Threshold View's missteps. Deontologists may worry that this approach will be too

---

[1] My efforts here are part of a broader project on deontological decision theory, see for example (Lazar 2015a, 2016, 2017b). While this paper was under review at *Philosophical Studies*, I was made aware of a paper by Horacio Spector (2016) which reaches similar conclusion to mine in this paper, albeit by a different route.

[2] For views more sympathetic to deontology, see Colyvan et al. (2010), Raz (2011), and Zimmerman (2008).

permissive, and that it is computationally over-demanding. I answer those objections.

I think the prospects are good for deontological decision theory. But this paper is just a start. I aim simply to persuade deontologists to abandon the Threshold View, and show them not only that we can use decision theory without sacrificing our deontological commitments, but that we have positive reason to do so.

2 Groundwork

First, though, some groundwork. This paper is about deontological ethics, subjective permissibility and the ethics of killing. I briefly explain what I mean by the first two terms, and offer an outline of a theory of the objective morality of intentional killing.

I lack a definitive account of what parts deontologists from consequentialists. Perhaps there is no sharp line—just clusters of loosely connected positions (Dougherty 2013a). At the abstract end, consequentialists exclusively derive deontic conclusions (about what we ought to do) from evaluations of states of affairs. Deontologists do not (Schroeder 2007; McNaughton and Rawling 1998; Scheffler 1994; Smith 2009). For deontologists but not consequentialists, the right is prior to the good (Rawls 1999).

Deontologists often think that some of our reasons are agent-relative—for example, that I have a special reason to ensure that I don't kill, not merely an agent-neutral reason to ensure that all people do not kill (Broome 1991; McNaughton and Rawling 1991).[3] Many deontologists think that one's mental states—for example, one's intentions or reasons for acting—are relevant to objective permissibility. Others think that the specific kinds of causal relationship between harms inflicted and goods achieved thereby are morally relevant (Kamm 2007). Most deontologists think that we have rights, which protect us against marginal interpersonal trade-offs. And they typically think that agents are sometimes permitted to realise suboptimal outcomes, just because the cost to them of bringing about the best outcomes is too great (Schroeder 2011). For a standard of subjective permissibility to work for deontological ethics, it must be able to capture the deontological stance on each of these issues, from the most abstract to the most substantive.

What, then, of 'subjective permissibility'? I define it in contrast with objective permissibility. By 'objective', I mean in light of all the relevant non-moral facts. By 'subjective', I mean in light of an agent's subjective epistemic circumstances. 'Subjective permissibility' is an umbrella for all non-objective permissibility standards. It could be indexed to the agent's beliefs, her evidence, or someone else's beliefs or evidence. These are important differences, but none of the problems discussed in this paper depends on a particular epistemic standard being in place. I also ignore moral uncertainty, which raises quite different issues (Ross 2006).[4]

---

[3] Another candidate is time-relativity (Broome 1991; Colyvan et al. 2010).

[4] I cannot defend my reasons for ignoring moral uncertainty in detail here, but, in broad strokes, they are mostly consistent with those set out in Hedden (2016).

Some think morality can support only one standard of rightness, so all-out objectivists and all-out subjectivists must fight it out (Graham 2010). Others are more tolerant of a 'proliferation of oughts'. I cannot weigh into that debate here. I concentrate instead on establishing what a standard of subjective permissibility should look like if we need one. Only all-out objectivists should baulk at this task, and even they need *some* account of what a morally conscientious person should do given her uncertainty.

But what do I mean by a standard, or criterion, of permissibility? Ideally, such a standard would give us a set of necessary and sufficient conditions for an act's being permissible. Developing and defending such a standard of subjective permissibility for deontological ethics is too ambitious a task for one paper. Instead, I aim only to rule out one proposed necessary condition for a lethal act's being subjectively permissible, and introduce an alternative sufficient condition for an act's being subjectively permissible.

A standard of subjective permissibility is just that: a *standard*. It is not a decision procedure. I will return to this point below, but I doubt whether moral philosophy alone can provide us with decision procedures that we can actually operationalise in real life. Decision procedures should be attuned to the particular psychology of the person applying them, as well as to the circumstances of her action. A standard of right action simply picks out the option that she should choose from any given option set. It does not tell her how to identify that option in the exigent circumstances that she actually faces.

It is also worth stressing that deontologists can accommodate non-moral uncertainty in other ways, besides developing a standard of subjective permissibility. They might, for example, propose a series of free-standing principles for dealing with risk.[5] Or they might advocate cultivating virtues of judgement and restraint, which allow us to negotiate our uncertainty. They might instead propose rules of thumb to which deontologists should adhere. I cannot, in this paper, knock down all these alternative approaches. But I submit that defending a criterion of subjective permissibility is at least one attractive candidate approach. Even if these other methods are also important, there should be room for a standard of subjective permissibility in any moral theory.

So much for deontological ethics and subjective permissibility. It remains to summarise the account of the ethics of killing that I am presupposing in this paper. According to most philosophers working in this area, people have rights to life, which can be overridden or lost only exceptionally.[6] A target who retains her right to life can be justifiably killed if her death achieves a good great enough to override that right. This is a *lesser-evil* justification for killing. Some think lesser-evil justifications can apply even when the good achieved is not significantly greater than the harm done, if some killings are less objectionable than others (for example, because merely foreseen, not intended), or if we invoke agent-relative reasons, such

---

[5] Subjectivist theories of rights fall in this category (McCarthy 1996, 1997; Zimmerman 2008; Quong 2015).

[6] This summary is intended to be general enough that most contemporary deontological theorists of killing can broadly endorse it.

as a permission to favour one's own interests or an associative duty to protect one's family (Lazar 2013).

We need to pause here, to make very clear a central failing of the existing literature on uncertainty and deontological ethics. A number of critics have argued that deontologists have no plausible way to extend their theories to decision-making under uncertainty (Jackson and Smith 2006, 2015; Huemer 2010); some deontologists have answered their criticisms (Hawley 2008; Aboodi et al. 2008). The whole debate has drawn heavily on examples from the ethics of killing. But the focus, in every case, has been on an *absolutist* approach to the ethics of killing, which brooks no trade-offs between the right to life and other considerations—or, in other words, denies the possibility of a lesser evil justification for intentionally killing the innocent. But I know of only one deontological philosopher working on the ethics of harm who professes to be an absolutist, and even he would not endorse this extreme version of absolutism.[7] Everyone else is a 'moderate deontologist', who believes that intentionally killing the innocent is sometimes permissible, if the good achieved thereby is great enough.[8] This endorsement of lesser evil justifications will be crucial in what follows.

But killing is not justified only when it is a lesser evil. It can also be justified if the target has lost the protection of his right to life.[9] Some think we can waive that right; some think those who commit heinous crimes deserve to be killed. And sometimes a target is liable to be killed in self- or other-defence.

N is liable to be killed to avert some unjustified threat if and only if: N is sufficiently responsible for that threat, and killing N is a necessary and proportionate means to avert that threat. Opinions differ on what counts as 'sufficiently responsible'. Killing is necessary, in this context, when there is no other less harmful means to avert the threat.[10] Killing is proportionate when the threat is serious enough for killing to be an appropriate response. This is what Jeff McMahan calls 'narrow' proportionality (McMahan 2009).

Although N's lacking the protection of his right to life is sufficient to ensure that *he is not wronged* by being killed, it does not ensure that killing him is objectively permissible. For there might be other adverse consequences of doing so, which rule it out. Suppose, for example, that killing N would also kill a large number of bystanders. Or suppose N, and only N, knows the cure for cancer. Killing in such cases might be 'widely disproportionate', to use McMahan's (2009) terminology.

Research on killing has focused on objective permissibility, even under uncertainty.[11] This is an oversight. Whether the target is liable, whether she is

---

[7] Alec Walen, personal communication.

[8] This is true, for example, of Christian Barry, Kai Draper, Cécile Fabre, Adil Haque, Jeff McMahan, Jonathan Quong, David Rodin, Victor Tadros, Suzanne Uniacke. I genuinely cannot think of a single author working actively in this area, besides Walen, who is an absolutist.

[9] Some think that the right to life can be weakened without being lost, and so there can be hybrid justifications for killing, comprising both lesser-evil and liability components (McMahan 2011, 2014; Bazargan 2014).

[10] It's worth stressing that this is how the necessity constraint should be understood for the objective permissibility of killing.

[11] For two exceptions, see (Haque 2012; Zimmerman 2008).

sufficiently responsible, whether harming her is necessary or proportionate, what would happen if one did nothing: each of these matters depends on facts about which we are typically uncertain.[12] This uncertainty is especially acute in armed conflicts, when the people doing the killing are removed from the justifying cause, and their victims invariably include both liable and innocent people. We fight only dubious battles. We should calibrate our theories of permissible killing accordingly.

3 The Threshold View

Adil Haque (2012) has argued that killing is subjectively permissible only if it is more likely than not that the target is liable to be killed. Jeff McMahan (1994) once advocated the biconditional version of this view. Kim Ferzan (2005), Helen Frowe (2010) and (later) McMahan (2011) have argued that an individual can make himself *objectively* liable to be killed by being responsible for another person's *justified belief* that he is objectively liable, where part of what makes a belief justified is that it crosses some probabilistic threshold. Michael Zimmerman (2008) thinks one can lose one's right to life merely because it is likely enough that one is liable (regardless of responsibility). In their critiques of absolutist moral theories, Michael Huemer (2010), Frank Jackson and Michael Smith (2006, 2015) have attributed to them a threshold-based view of this kind. Hawley and Aboodi et al. (2008) have defended it against those criticisms. Outside philosophy, we find echoes of this view in the standard of proof in the criminal law—'beyond a reasonable doubt' (Walen 2015). And the Obama administration professed to permit targeted killing only if there is 'near certainty that the terrorist target is present' and 'near certainty that non-combatants will not be injured or killed' (Office 2013).

The central element in each of these views is this: killing is subjectively permissible only if it is sufficiently likely that the target is liable to be killed. But recall that on the kind of deontological view being considered, to which all of the philosophers just mentioned subscribe, intentionally killing the innocent can be justified as a lesser evil even when the target is not liable to be killed. The full statement of their view must then be something like this:

> *The Threshold View*: Intentional killing is subjectively permissible only if either it is sufficiently likely that the target is liable to be killed, or the subjective lesser evil condition is triggered.[13]

---

[12] Note that although each of these questions is of course a moral question, they each depend on non-moral facts, so we are still in the realm of non-moral uncertainty. Of course, we might also be uncertain about which standard of liability (for example) to apply. But as already noted, I am setting aside moral uncertainty in this paper.

[13] Thanks to a referee for pressing me on the best formulation of the Threshold View. I have agonised over this point. There are potentially many different threshold views—for example, one could parse out all the elements of one's objective theory of the permissibility of killing, and apply a threshold to each of them. One obvious choice: intentional killing is subjectively permissible only if either (i) it is sufficiently likely your target is objectively liable, that killing her is objectively necessary, and that killing her is objectively proportionate or (ii) it is sufficiently likely that killing your target will realise enough objective good to be objectively permissible as a lesser evil. Alternatively one could use thresholds only

I do not think anybody has yet explored what the 'subjective lesser evil condition' would be. There seem to be at least two possibilities: first, the target is not sufficiently likely to be liable, but the expected good achieved by killing her is sufficiently great; second, the target is not sufficiently likely to be liable, but it is sufficiently likely that the good achieved will be great enough to justify killing an innocent person.[14] Both approaches are in the spirit of threshold deontology: one applies the threshold to the amount of expected good achieved, the other to the probability. Since I will focus on the liability component of the Threshold View, I will remain neutral between these possibilities.

As already noted, the Threshold View gives a necessary condition for intentional killing being subjectively permissible. To develop it into a standard of subjective permissibility for intentional killing, we would need not only to flesh out the subjective lesser evil condition, but also to incorporate something like the wide proportionality condition: an act of intentional killing can clearly be subjectively impermissible in virtue of its consequences for those who are not its target, if bystanders are collaterally harmed, for example. Again, we could interpret this in terms of expected goods, or in terms of it being sufficiently likely that the act is widely proportionate. Indeed, we could ultimately go the whole way and develop a criterion of subjective permissibility according to which an act (any act, not just one of intentional killing) is subjectively permissible if and only if it is sufficiently likely to be objectively permissible. However, there are many decision points on the route to that destination, and I do not want to saddle deontologists with a theory that they have not explicitly defended. With that in mind, I concentrate only on the Threshold View as stated, and in particular on its first clause.

To better understand that clause, we must first ask what 'sufficiently likely' means. To keep things simple, let's say the probability that the individual is liable must be at least 90%. These probabilities must be subjective in some sense, but I will not distinguish between actual credences, rational credences, evidential probabilities, and so on. This is an important topic, but my critique of the Threshold View does not depend on it. I also assume the threshold is fixed, since if the threshold varies with the stakes (a plausible idea) then the Threshold View will be a notational variant on the decision-theoretic approach proposed in Sect. 4 (Lazar and Lee-Stronach 2016).

Some might quibble about the arbitrariness of choosing a number for a probability threshold. They might also dispute the plausibility of there being a sharp cut-off, so that if you are 89.999% confident, it's a no-go—but once you get that additional 0.001% confidence, it's subjectively permissible. These would be cheap shots.[15] We should take the implied precision of numbers with a pinch of salt. Even if our moral reasons are not metaphysically indeterminate, they are at least

---

Footnote 13 continued

for some of those conditions, or fold some of those conditions into the liability standard, and no doubt pursue other possibilities as well. However, I do not want to be accused of creating a straw man to knock down, so I have focused on that element of the Threshold View that others have defended in print.

[14] We can assume that 'expected good' is calculated in the usual way, by weighting the good achieved in the different possible outcomes of the act by the likelihood that each will come about.

[15] Jackson and Smith (2006) agree that this is not a serious worry.

epistemically vague (Dougherty 2013c). We cannot pin them down to infinite decimal places.

Applied to the ethics of killing, the Threshold View is initially quite appealing: killing is a serious business, and with a high threshold, the Threshold View reflects an appropriately cautious approach to risking killing the innocent. It also promises to give usable advice. I will, however, lay two charges against the Threshold View. First, it declares subjectively impermissible some objectively unnecessary acts of self-defence that are intuitively subjectively permissible. Second, it adopts an inconsistent attitude towards trade-offs between risks of killing an innocent person, and the goods that can be achieved thereby, depending on whether those trade-offs have to do with objective or subjective permissibility.

First: many philosophers of self-defence think that 'necessity is internal to liability' (McMahan 2009). That is, one cannot be liable to suffer a harm that is unnecessary to avert a threat. If this is right, then The Threshold View must be wrong, as the following case shows:

> *RPG*: Ruth culpably threatens Bruce's life. He has time to fire one weapon at her before she kills him. Before him are a rocket-propelled grenade (RPG) and two pistols. With a loaded pistol Bruce, a crack shot, could halt Ruth's attack without killing her. The RPG will certainly kill her. Bruce knows Ruth is culpable, and knows one of the pistols is loaded, but he does not know which one.[16]

In this case, Bruce knows for sure that Ruth is not objectively liable to be killed, because he knows that killing her is objectively unnecessary, since he could pick up the loaded gun and merely wound her. And yet, intuitively, it is highly plausible that it is subjectively permissible for him to kill Ruth. Wounding Ruth and surviving would be better than killing her, but Bruce need not run a 50% risk of death to save a culpable attacker the difference between being killed and being wounded. The Threshold View's exclusive focus on likelihood of objective liability obscures how bad things will be for Bruce if he picks the wrong gun. Given that risk, it seems subjectively permissible for him to fire the RPG, even though he knows Ruth is not objectively liable to be killed.

Interestingly, this example also decisively undermines an approach to subjective permissibility that McMahan and others sometimes espouse. McMahan asserts that 'an act is subjectively permissible or justified when two conditions are satisfied: first, the agent acts on the basis of beliefs, or perhaps reasonable or justified beliefs, that are false, and, second, the act *would* be objectively permissible or justified if those beliefs were true'.[17] Now, set aside the fact that this definition works only for acts that are *merely* subjectively permissible, whereas an act could be both subjectively and objectively permissible. More seriously, RPG shows this account of subjective permissibility to be wrong. It is clearly subjectively permissible for

---

[16] This is an application to the ethics of self-defence of a familiar conundrum, first presented in (Regan 1982), and later developed in (Jackson 1991; Parfit 2011).

[17] For example, McMahan (2009).

Bruce to kill Ruth in self-defence. But he rightly believes that killing her is objectively impermissible, since if he knew all the facts, he would know which gun is loaded, and be able to stop her attack without wounding her.

Of course, not all theorists of self-defence believe that necessity is internal to liability (Frowe 2014; Firth and Quong 2012). Perhaps some of them will view RPG as evidence in favour of their view. And yet, everyone agrees that satisfying the necessity condition is a prerequisite for intentional killing being objectively permissible. So any theory of the subjective permissibility of killing must also incorporate some version of the necessity condition. And if that makes use of a threshold—for example, by adding the clause that it must be sufficiently likely that killing one's target is objectively necessary—then RPG continues to apply in full force. Conversely, if the necessity constraint does not use a threshold, as for example in the subjectivist necessity standard that I defend elsewhere, then the effect is to build something very close to the decision-theoretic approach directly into our decision rule (Lazar 2012).[18] That would be a pyrrhic victory for fans of thresholds.

There is a second, more fundamental problem with the Threshold View. The central feature of the Threshold View is that we must cross the same high threshold regardless of what the stakes are. So, consider the following two cases:

> *High-Value Target*: A CIA pilot has a man in the sights of her predator drone. She is hunting the second-in-command of ISIL. Killing him is likely to save many lives. She is not sure that the person in her sights is her man.

> *Low-Value Target*: Same as *High-Value Target*, but this time she is hunting a low-level ISIL operative. Killing him will save fewer lives. She is again uncertain whether this is her man, to just the same degree as in *High-Value Target*.

The Threshold View says that the same threshold applies in these two cases. The better expected results in the first case are irrelevant: the threshold stands firm, prohibiting any trade-off between the additional risk of killing an innocent person and the additional good achieved by killing the high-value target. But deontologists like those described in Sect. 2 already accept these kinds of trade-offs. That is, they already believe that we may override people's rights to life when doing so achieves a sufficiently great good. In fact, trade-offs arise not only in lesser-evil justifications, but also within the liability justification, which is inherently instrumental: whether you are liable to be killed depends in part on how much good is achieved by killing you (this is captured by the narrow proportionality constraint).[19] All those who have engaged in this debate thus far have missed this point, because they all focus on absolutist deontological theories, which do not brook trade-offs (Isaacs 2014; Jackson and Smith 2006, 2015; Aboodi et al. 2008; Hawley 2008). But, as already noted, almost nobody working on the ethics of harm is an absolutist in that way. And if we accept trade-offs between goods achieved and harms to the innocent in

---

[18] Incidentally, though it is a bit awkward—to say the least!—a subjectivist necessity constraint could still be internal to an objective liability standard. Thanks to a referee for pressing me on this point.

[19] See especially McMahan (2009).

our theory of the objective permissibility of killing, then how can we deny them in our theory of subjective permissibility? This is the fundamental inconsistency at the heart of the Threshold View.

One might respond as follows: if our objective moral theory has thresholds, then shouldn't our subjective moral theory do so as well?[20] But I think that conflates two distinct kinds of thresholds. When we weigh one consideration against another, the tipping point is a kind of threshold, at which one outweighs the other. But this threshold is obviously totally consistent with weighing the goods that we can achieve against the wrongs that we commit, that is, with giving due consideration to all the reasons that might bear on your action. In objective deontological moral theory, thresholds are primarily *permissive* concepts: if the good achieved is above some threshold, then it is permissible (for example) to infringe a right in its pursuit.

The epistemic threshold employed by the Threshold View is quite different: instead of identifying a tipping point at which one set of reasons outweigh another, it screens off reasons from consideration, without regard to their weight or magnitude. Instead of a permissive threshold, it is a *restrictive* one: it says that no amount of expected good, however great, can justify running an additional risk of a given magnitude of killing an innocent person. It posits a discontinuity between the considerations at stake, where objective thresholds do just the opposite.

The Threshold View asserts that some degree of expected good can justify actions that have a 9% probability of intentionally killing an innocent person, but *no* additional expected good, no matter how great, can justify running an additional 2% risk of killing an innocent person. This is a species of absolutism, which, as I have emphasised, is endorsed by very few practising deontologists. If we think that some actual goods are valuable enough to justify knowingly intentionally killing the innocent as a lesser evil, then we *must* think that some expected goods can justify running an additional 2% risk of intentionally killing an innocent person.

One might respond that this is where the lesser evil part of the Threshold View kicks in: if the expected good is great enough, then this additional risk would be justified on the same grounds that it is permissible to knowingly kill an innocent person for the sake of a greater good. That might preserve some intuitive plausibility for the Threshold View in some cases, but it would come at the cost of positing that there is no difference between the amount of expected good needed to justify running an 11% risk of your target being an innocent person, and that required to justify killing your target knowing that he is innocent. That is a deeply implausible result, and still involves a groundless inconsistency in the treatment of trade-offs under risk and under certainty.

We can expose this inconsistency in a number of ways. First, we can turn risky cases into lesser-evil cases, by repeating them enough times (Dougherty 2013b; Norcross 1997). If the pilot pulls the trigger each time she faces a situation like *High-Value Target*, then she will kill more innocent people, but also save many more lives than if she does so each time she faces one like *Low-Value Target*. Suppose, for example, that the odds are such that over a hundred iterations the pilot in

---

[20] Thanks to a referee for this response.

High Value Target is likely to kill, on average, two innocent persons, while the pilot in *Low Value Target* is likely to kill, on average, one innocent person. And suppose that the good achieved over those hundred iterations of *High Value Target* would be enough to justify killing two innocent people as a lesser evil, in a single case, while the good achieved by the pilot of *Low Value Target* over those hundred iterations would not be enough to objectively justify killing one person as a lesser evil in a single case. These outcomes are perfectly consistent with standard deontological ethics. But if the goods achieved can objectively justify the innocent lives taken when considered over multiple iterations, why can't they render firing in a single case subjectively permissible, when we are trading off risks instead of actual deaths? Why should we take one attitude to trade-offs when considering campaigns, but a completely different one when thinking about individual actions in isolation?[21]

Or consider a second group of cases: in the first, Bruce is 90% confident that Cody is liable to be killed, and sure that killing him will save his own life; in the second, Bruce is 85% confident that Cody is liable to be killed, and sure that killing him will save his own life, and the lives of 20 other people; in the third, Bruce is sure that Cody is not liable, but killing him will save 1000 people (or however many is required to make killing Cody a lesser evil). According to the Threshold View, Bruce's action is subjectively permissible at the bar of liability in the first case, and permissible at the lesser evil bar in the third case, but is impermissible in the second case. But if we are prepared to trade off certainly killing an innocent person against saving some number of lives, then we should be prepared to trade off an additional marginal risk of killing an innocent person against some additional number of lives saved. Perhaps the ratio is steep—indeed, as I argue below, I think that the more likely it is that Cody is innocent, the steeper the increase in the number of lives that must be saved for killing him to be subjectively permissible. But the Threshold View occupies an uneasy middle ground: it endorses lesser evil justifications for certainly killing an innocent person, but denies their role in justifying additional marginal risks of killing the innocent.

If we are prepared to trade off actually killing nonliable people against goods achieved, we must also be prepared to trade off risks of killing nonliable people against prospects of achieving greater goods. The Threshold View prohibits the second kind of trade-off, but its advocates endorse the first. That balancing act seems impossible to pull off. On, now, to another one.

4 Deontologists can maximise expected moral utility

4.1 Introducing the decision-theoretic alternative

I think deontologists should reject the Threshold View, and instead adopt a broadly decision-theoretic standard of subjective permissibility. This is its central claim:

---

[21] No doubt there are some complex issues here, which I address in detail in (Lazar and Lee-Stronach 2016). But the basic point is quite simple.

The Decision-Theoretic Standard: An act is subjectively permissible if it maximises expected moral utility.

In this section, I first outline the Decision-Theoretic Standard, and answer a preliminary, and obvious, worry that this standard is ill-suited to a deontological approach to ethics. I then show how the Decision-Theoretic Standard affords a better approach to the ethics of killing than does the Threshold View, before answering two further objections—that the Decision-Theoretic Standard will be too permissive, and that it is computationally overdemanding.

Invoking decision theory immediately invites the question: *which* decision theory (Buchak 2016)? There are many possibilities; decision theory is no more settled a field than is ethics. I endorse the naïve, or (less pejoratively) classical, approach to decision theory, which takes the agent's probability and moral utility functions to be independently established, then tells the agent to maximise expected moral utility.[22] This contrasts with the dominant approach in the twentieth century, which takes agents' preferences over gambles as given, then argues that if those preferences obey some reasonable-seeming axioms, the agent can be represented as an expected moral utility maximiser, with a probability function and a moral utility function that are unique up to positive linear transformation (Von Neumann and Morgenstern 1944). Interestingly, the naïve approach has recently seen a resurgence, following criticisms of those representation theorems and the uses to which they have been put (Easwaran 2014; Meacham and Weisberg 2010; Peterson 2002). Regardless of decision-theoretic fashions, however, the naïve approach is the only plausible candidate for our purposes. Whatever the theoretical merits of arguing that adherents to any moral theory can be represented as expected utility maximisers,[23] those arguments are of little practical use, since they work only if we can input the agents' preferences over gambles—but our goal is precisely to find out what preferences over (moral) gambles deontological agents should have.

The Decision-Theoretic Standard has acts, states and outcomes as its moving parts.[24] If an agent performs act A, given that the world is in state S, outcome O will result. Acts are options, things the agent can do.[25] States are mutually exclusive and jointly exhaustive ways the world might be. Outcomes are the worlds that result from an act, given some state.

Having identified the relevant acts, states and outcomes, the agent's next step is to rank the outcomes. This means identifying the properties of each outcome that count in its favour and against, and bringing them together in a summary verdict of its objective choiceworthiness. She must rank outcomes on an interval scale: ordinality will not do. She must indicate how strong her overall reasons are to avoid or to seek each outcome, according to her objective moral theory. We can represent

---

[22] See, for example, Pascal's Wager, in the Pensées, part III, Section 233.

[23] For this approach, see: (Oddie and Milne 1991; Dreier 2004, 2011; Colyvan et al. 2010).

[24] For simplicity I use terminology familiar from Savage's (1954) decision theory, but do not thereby pin my flag to that mast.

[25] Each of these concepts undoubtedly raises innumerable complexities, but I doubt whether any will be relevant to my practical purposes in this paper. So I ignore them with a clear conscience!

this ranking with numbers, which together constitute the agent's moral utility function. Because of vagueness, or at least imprecision, the ordering might not be complete.

Obviously, though, whether she should choose an act depends not only on those rankings, but also on how likely the outcomes are to be realised. This depends on what state the world is in. So the agent needs a probability function as well, which assigns probabilities to states given that one acts a particular way. I will not stipulate which probabilities matter: as noted, I am focusing on subjective rightness in general. I also ignore vague credences and radical uncertainty, in which the agent lacks a probability function. These are merely attempts to keep things simple, not statements of allegiance.

Next, the agent should combine her probability and moral utility functions, weighting the moral utility of each outcome by the probability of the state that leads to it. The sum of the products of these numbers for all an act's outcomes is its expected moral utility. An act is subjectively permissible if its expected moral utility is no lower than any other's.

4.2 Decision theory is for deontologists too

So far, so familiar—and so unappealing to deontologists. This simply looks like endorsing a consequentialist subjective permissibility standard. But appearances are misleading: adopting the Decision-Theoretic Standard is consistent with respecting both the structural and substantive commitments that divide consequentialists and deontologists.

First, we can adopt that standard without affirming that what one ought to do is determined by evaluations of states of affairs. What we subjectively ought to do is determined by our rankings of outcomes (combined with our probability function). But how we rank outcomes is itself determined by our objective moral reasons, which govern what we objectively ought to do. When ranking an outcome O of an act ϕ, we need to work out what our objective moral theory says about realising O given that you ϕ. If you ϕ, and O comes about, then have you acted objectively permissibly? If so, what are the objective moral reasons in that act's favour, and how weighty are they? If not, then how seriously wrong was ϕing? The order of explanation runs like this: the objective right determines the objective 'good', the probability-weighted expectation of which determines the subjective right. So, the right remains prior to the good.[26]

Of course, this means that our objective moral theory must tell us more than just whether some act is objectively permissible or impermissible: we also need to

---

[26] Contra the concluding remarks in Oddie and Milne (1991). In a paper that responds, in part, to this one, Douglas Portmore (2016). argues that the existence of fundamental indeterminacy might ultimately undermine this claim. Though it seems unlikely that the outcome of the most fundamental debate in normative ethics could turn on metaphysical claims about indeterminacy, I agree with Portmore that deontologists need a response to his arguments, but will not broach that topic in this paper. My concern here is much more pragmatic: I want a criterion of subjective permissibility that works for people with the kind of deontological view described in Sect. 2. More fundamental questions about what *ultimately* grounds our moral reasons are, I think, inevitably intractable, as well as beyond the scope of my enquiry.

know the extent to which our objective reasons tell for or against our action. But that is no problem: it is surely obvious that even if stealing a loaf of bread and killing a prison guard are both impermissible, one is more seriously wrong than the other. And the objective moral reasons in favour, for example, of saving 100 lives are weightier than those in favour of saving 50 lives.

It is important to stress that moral utility is just a formal construct. It does not describe a property, *value*, or *goodness*, which is maximised in the highest-ranked outcomes. We might equally call utility 'choiceworthiness', as I do elsewhere (Lazar 2017b). It certainly doesn't correspond to anything like well-being or preference-satisfaction. It is merely a way of representing a ranking, which is grounded in our objective moral theory. If our objective moral theory (bizarrely) ranked lowest the outcomes in which, say, happiness is maximised, then they would be the ones with the least moral utility. Decision theorists will of course be familiar with this point. But deontologists typically worry that 'play[ing] the consequentialist value-game' is constitutionally unavailable to them (Aboodi et al. 2008). That worry is mistaken.

Moral utility rankings should not scare deontologists. Nor should the focus on outcomes. As has long been recognised in decision theory, an act's outcome can include everything that matters about the world realised by an act (Broome 1991; Jeffrey 1965; Dreier 1993). This includes facts about the world's history, its having been brought about by that act, the agent's intentions, beliefs and so on, its relationship to non-actual worlds, agent-relative reasons as well as agent-neutral ones, the option set from which that act was chosen—indeed, anything your objective moral theory deems important. All we need do, when ranking each outcome, is consider what our objective moral theory would say about φing, given that O results.

As for the more substantive commitments of deontologists, with two exceptions, I think these can all be captured within the deontological moral utility function.[27] That is, they can all be translated into properties of outcomes, which we then rank according to the objective permissibility of the act, given that it results in that outcome. This much is familiar from the literature on 'consequentialising' moral theories (Dreier 1993; Louise 2004; Peterson 2010; Portmore 2009). Agent-relative reasons, the significance of intentions and causal structures to permissibility, the role of rights: all of these can be incorporated into our ranking of outcomes. Most pertinently, the moral significance of the contrast between harming someone who is and is not liable to be killed can readily be accommodated within the moral utility function.

At this point deontologists will complain (and consequentialists will crow) that if we can really represent deontological theories in this way, then we're all consequentialists after all (Oddie and Milne 1991; Dreier 1993; Smith 2007). Both sides are mistaken. The task of finding a deontological criterion of subjective permissibility both vindicates consequentialising and puts it in its place. It shows that

---

[27] The first exception is discussed below; the second has to do with deontological opposition to aggregation. I think it's quite plausible that deontologists should resist some kinds of aggregation—for example, when we must choose between saving one person's life, or sparing some much larger number a much more trivial harm. I show elsewhere (Lazar 2017a) that representing this kind of view within a decision-theoretic framework is extremely difficult to do.

even if all the many criticisms of consequentialising are true—that it at best delivers the right verdicts on cases, without having any explanatory power, that it obscures what motivates moral agents, that it is 'gimmicky', that we can also 'deontologise' any moral theory, and that the 'compelling idea' that it is always permissible to bring about the best outcome is not that compelling after all—there is still a reason to consequentialise our deontological moral theory: because doing so allows us to use the tools of decision theory for moral decision-making under uncertainty.[28]

So: that we can represent all objective moral theories as consequentialist does not foreclose any important moral debates. All the objections against consequentialising might be right. Far from being the last nail in deontology's coffin, the possibility of representing deontological moral theories in this way saves them from a much more threatening objection: that they lack a criterion of subjective permissibility. But at the same time, the problem of non-moral uncertainty provides an answer to the most pressing objection against consequentialising—that it is a pointless exercise (Tenenbaum 2014; Dreier 2011). To the question: why bother? We can now answer: because it renders our deontological moral theories amenable to incorporation within a decision-theoretic approach to moral decision-making under uncertainty, which is no small thing.[29]

It is important to stress, however, that the Decision-Theoretic Standard is only a sufficient condition for subjective permissibility. Deontologists typically think that we are sometimes permitted to realise options that are morally suboptimal. In other words, an agent cannot be *required* to maximise expected moral utility, when doing so involves her bearing an unreasonable expected cost. These kinds of moral options raise many interesting and complicated issues, which I discuss in a separate paper (Lazar 2017b), although I return to this point briefly below.

4.3 Decision-theoretic killing

I'm selling the Decision-Theoretic Standard as a general approach to subjective permissibility for deontological ethics. To make the case in its favour, I need to show that it handles the ethics of killing better than the Threshold View does, by illustrating how it applies to the cases introduced in Sect. 3.

The Threshold View says that it is impermissible for Bruce to kill Ruth in RPG. The Decision-Theoretic Standard could deny this on two independent grounds. Given Ruth's culpability, it is highly likely that killing her maximises expected moral utility. Even if killing her is not objectively necessary, it seems clear that it is morally preferable to the innocent party having to bear a 50% risk of being killed, because we must weight the interests of a culpable attacker considerably less than we weight those of an innocent defender.[30] But even if this is wrong, then the

---

[28] For those criticisms, see especially (Schroeder 2007; Tenenbaum 2014; Hurley 2013; Vallentyne 1988).

[29] Thus my argument is part of what Andrew Schroeder (Forthcoming) would call the 'pragmatic' case for consequentialising.

[30] This move is unavailable to adherents to the Threshold View. If they think necessity is internal to liability, then it is not sufficiently likely that Ruth is liable (even if it is highly likely, indeed known, that

additional conditions that flow from a natural extension of the Decision-Theoretic Standard to accommodate the option to act suboptimally will make room for Bruce's action to be permissible: he cannot reasonably be expected to bear a 50% risk of being killed in order to spare the life of someone who is culpably attacking him (Lazar 2017b).

The Decision-Theoretic Standard also allows us to adopt a consistent approach to trade-offs. Just as we accept lesser-evil justifications in decision-making under certainty, we also accept trade-offs between risks and opportunities (where opportunities are probabilities of good outcomes). Just as it might be permissible, over a series of trials, to kill more innocent people in the pursuit of a high-value terrorist target than in pursuit of a low-value one, so can it be subjectively permissible to proceed against a high-value target when doing so runs a greater risk of killing an innocent person than could permissibly be run to secure a lower-value target. Similarly, the additional 20 lives that can be saved in the final case of Sect. 3 can be relevant to whether it is permissible to run an additional 5% risk of killing an innocent person.

Ultimately, the Decision-Theoretic Standard is just a formal framework: its substantive verdicts will depend on how we flesh out the moral utility function, and on the details of each particular case. But the basic difference from the Threshold View is very clear. The latter permits us to weigh goods achieved against wrongs inflicted at the objective level, but screens them off at the subjective level. The Decision-Theoretic Standard permits trade-offs at both levels.

Here is a more general way to see the difference between these views. We can distinguish between the *grounds* of a verdict of objective permissibility—or proportionality, liability, necessity, lesser evil, and so on—and the verdict itself. The Threshold View cares only about how close one's action is to being objectively endorsed in each of these categories. The Decision-Theoretic Standard cares about the *grounds* of objective permissibility, and weights each for its likelihood of being at stake. In other words, the Decision-Theoretic Standard invites us to consider, directly, people's interests, their rights, and the many other objective moral reasons that might be at stake in our action. The Threshold View instead invites us to consider each possible outcome of your action, and assess whether it satisfies those objective criteria. At issue between these approaches, then, is whether we ultimately care about these objective verdicts, or about the moral reasons that ground them. I think the latter makes more sense.

To illustrate this point, consider a possible addendum to the Threshold View. Its adherents have to find some way of accommodating a subjectivised version of the wide proportionality constraint—corresponding to the objective wide proportionality constraint, which says (roughly) that an act of intentional killing is permissible only if the evil unintentionally inflicted is justified by the good achieved. A natural approach would be to say that an act is subjectively permissible only if it is

---

Footnote 30 continued

she is culpable). If necessity is external to liability, but still a necessary condition for permissibility, then it is still subjectively impermissible to proceed because it is not sufficiently likely to be necessary, again, regardless of whether we know that Ruth is liable. Thanks to a referee for pressing for clarification here.

sufficiently likely to be objectively (widely) proportionate. Again, set the threshold at 90%. Suppose that you can save five lives in one of two different ways. The first has a 90% probability of killing only one person as an unintended side-effect, but a 10% probability of killing 1000. The second has an 89% probability of killing only one person as an unintended side-effect, but an 11% probability of killing two people.

If what we care about is proximity to the objective verdict, then we should prefer the first act to the second, because it is sufficiently likely to be objectively proportionate. But if what we care about ultimately is the lives saved, as well as the lives taken, then we should prefer the second act to the first, because its downside is so much more tolerable. I think that verdicts of objective permissibility are grounded in the moral reasons that apply to our action, and so when thinking about subjective permissibility we should focus on those reasons, and weight them directly for their probability of being actual, rather than fetishize the objective verdicts.[31]

4.4 Is the decision-theoretic standard too permissive?

The Decision-Theoretic Standard is much more flexible than the Threshold View. It can be adapted to accommodate the subjective permissibility of using lethal force that one knows to be objectively unnecessary. And it does not have to adopt inconsistent attitudes towards trade-offs between risks of violating rights, and the goods that can be achieved thereby. And though deontologists are understandably wary of 'playing the value game' (Aboodi et al. 2008), they need not fear decision theory—it is capacious enough for them too. But while we're in the business of making trade-offs, does the Decision-Theoretic Standard have any vices that weigh against these virtues? The two most natural worries are that the standard will be too permissive and that it rests on an arbitrary mathematisation of morality, which is no easier to apply in practice than is a criterion of objective permissibility. I consider each objection in turn.

Recall *High-Value Target*, above. Suppose that the pilot can save *many* lives by killing her intended target. Is it really plausible that if the stakes were high enough, she could be subjectively permitted to fire even if she's more confident than not that this isn't her man? Even if it's one in a hundred?

This is a genuine problem, but I have already hinted at the solution. The decision-theoretic framework is a shell which we can fill out with our substantive commitments. Though it in principle allows trade-offs between any kinds of considerations, if we want to reflect hesitancy about some of those trade-offs, we can do so. In *High-Value Target*, as the probability that this person is her target declines, and the probability that killing him will be useless rises, the pilot's killing him gets progressively more seriously objectively wrong. The greater the risk she takes with his rights, the more seriously she wrongs him if he turns out to be innocent, and the lower those outcomes in which she takes this risk should be

---

[31] Much more could be said here, of course. For a valiant attempt to operationalise a subjectivist version of the wide proportionality constraint, in the spirit of the thresholders, see Patrick Tomlin's MS, 'Proportionality and Probability'.

ranked, relative to those in which she doesn't (Lazar 2015b, 2017c). The marginal moral seriousness of killing an innocent person increases as the probability that this person was innocent gets higher. This further counterbalances the rise in good done, so that as the probability that this is her target declines, the amount of lives at stake has to increase sharply to counterbalance the additional wrong done to her victim if he is innocent.

Though I defend this view at length elsewhere, I can offer a further argument that *something* like it must be true, at least if we are to represent the kinds of theories of the ethics of killing that most deontological ethicists endorse, alongside a plausible standard for when it is subjectively permissible to use lethal force in one-to-one cases. I submit that, when paired with such a standard, one cannot comfortably hold the following three views at once:

1. Some violations of the right to life are very stringently prohibited; they are objectively permissible as a lesser evil only if an extreme and unusual good is thereby achieved.
2. An agent's mental states—her reasons for acting, her beliefs, her intentions—cannot affect the objective permissibility of her actions.
3. In one-to-one cases of self-defence, it is subjectively permissible to use lethal force when one is reasonably confident (say, 70–80%) that one's target is liable to be killed.

Consider Alice and Bill, her business competitor. Alice can kill Bill and use his body to protect her against the threat posed by a hitman, whom Bill may have contracted. If she does nothing, she will die. If Bill ordered the hit, then killing him is objectively permissible. If he did not, then it is not.

Killing an innocent person to use his body as a shield is surely among the most serious rights-violations. According to 1, this could be permissible only if an extreme and unusual good were achieved thereby. Let us stipulate that saving 100 innocent lives is a significant enough good (you can adjust the number to match your intuitions).

If 2 is true, then if Alice kills Bill, who is innocent, her belief that he ordered the hit is irrelevant to the permissibility of her action. So her violation of his right to life is among the most serious, and could be permissible only if it achieved an extreme and unusual good, like saving 100 innocent lives.

Imagine that Alice's case is replicated 100 times over.[32] At the first iteration, she must decide whether to kill $Bill_1$, thus saving $Alice_1$, $Bill_2$ saving $Alice_2$, and so on. It would be objectively permissible to proceed only if no more than one of her victims is innocent, for then she would save 100 lives at the expense of one right-violating killing.

Now translate this back into the one-shot version of the case. Alice is subjectively permitted to kill Bill only if she is at least 99% confident that he ordered the

---

[32] This is simply intended to illustrate the point that if we view killing innocent Bill as the most stringently prohibited kind of killing, then we will be forced to adopt a very high threshold of confidence that he is liable in order for killing him to be subjectively permissible. Obviously this kind of repetition would raise some other issues as well, but they would not undermine this general point.

assassination. This is an extremely exacting epistemic standard! I doubt whether it is ever met in real life. It is much higher than the standard that 3 suggests is intuitively reasonable.

Deontologists cannot reject 1. Rejecting 3 means that too many real cases of self-defence will be subjectively impermissible. So we must reject 2, and identify a respect in which Alice's mental states when she kills Bill mean that, even if he turns out to be innocent, this is *not* the most serious kind of right violation—what is most stringently prohibited is killing someone in this way whom you *know* to be innocent. My account of the connection between probability and objective wrongdoing can vindicate this judgement; others might do so as well. It is worse to kill an innocent person when the probability that person was innocent was higher.

The worry that the Decision-Theoretic Standard is too permissive is misguided, but considering it has revealed another virtue of this approach. If we have any robust intuitions about subjective permissibility (such as that Alice doesn't need to be 99% confident that Bill ordered the hit for killing him to be subjectively permissible), then we can infer back from them to judgements about objective moral theories. We effectively have another lever with which to pump our intuitions in the pursuit of reflective equilibrium.

4.5 Who brings a calculator to a gunfight?

The final objection targets the mathematisation of our moral reasons required by the decision-theoretic approach. Where do these numbers come from? How can we think that what one morally ought to do can depend on mathematical functions? How can we possibly expect anybody facing an actual decision problem— especially when lives are at stake—to whip out their calculator and work out the expected moral utilities?[33]

I have three responses, besides the caution that there is simply no better alternative available at present. The first is to insist on a degree of theoretical distance from the formal posits of decision theory. The Decision-Theoretic Standard gives us a framework in which to model moral decision-making under uncertainty. That framework uses representational devices such as numbers to help us decide which risks to avoid and which to take. Those devices are unavoidably imperfect. They carry implications which we may need to repudiate. In particular, the use of numbers implies that what they represent is precise, additive and multipliable. But we should ignore these implications unless they are grounded in our moral theory. Precision is an artefact of the model. It would be absurdly hubristic to think that we could pinpoint our moral reasons with any precision; indeed, they may well be fundamentally indeterminate. We could use a complex mathematical device to represent imprecision, for example sets of numbers, but even that would take the model too literally (Cook 2002). Better to just recognise that the implied precision is an artefact, and allow, for example, that if an option's expected moral utility is very close to that of the optimal option, both are subjectively permissible.

---

[33] These worries abound within decision theory too.

Nor should we be misled by the appearance of additivity and multiplicability. When settling on a ranking, we consider facts about each outcome and weigh up the reasons they generate, but how we reach an overall ranking depends on our objective moral theory. We *could* simply add up all the relevant reasons, assuming they combine additively. But we can also reach an overall verdict by considering each outcome holistically. Indeed, it could be entirely particularistic. The moral utility function needs only the summary verdict.

Some might still worry about the arbitrariness of fitting a number to our moral reasons. I think most of those worries are primarily about the choice of unit and origin. But the Decision-Theoretic Standard needs only an interval scale, with a sense of the relative distance between the outcomes at stake. And those comparative judgements do seem more tractable, at least in some cases. Moreover, balancing of this kind is an ineliminable part of our objective moral theories as well—for deontologists too, given both the use of lesser-evil justifications and the inherently instrumental nature of liability. Of course, precision is unattainable, but if we simply throw up our hands and deny any possibility of assigning weights to our reasons, we leave our moral theories utterly ill-equipped to deal with the practical dilemmas that we inevitably face.

What about the objection that no real person could, or would, ever reason like this? One initial motivation for adopting a standard of subjective permissibility was that theories of objective permissibility offer us useless advice when we do not know the facts on which our actions' objective permissibility depends. But identifying the relevant outcomes, assigning probabilities to the states, formulating a moral utility function: none of this is easy (Feldman 2006). Sometimes it involves complex maths; in general it requires time, patience and a calculator. People whose options include killing typically lack these resources. Telling them to maximise expected moral utility might be as useless as telling them to do what is objectively right.

Though full calculations of expected moral utility might be too technically complex, decision theorists have already started developing a 'realistic decision theory', from which ethicists can learn (Weirich 2004). We can develop simpler versions of decision-theoretic calculations that are tractable for ordinary decision-makers, and yet will typically deliver the same results as more complex calculations (Kahneman et al. 1982; Lin 2013). For example, we often need not enumerate every possible outcome of our acts, but can get the same results by concentrating only on the most important ones; similarly, we can go a long way with only a partial ordering of outcomes; and qualitative measures of probability and moral utility can often deliver the same results as quantitative ones (witness standard risk-assessment matrices, for example). The Decision-Theoretic Standard might not give us advice that we can use in every situation, but it sets a benchmark that we can profitably aim at, which is more than we can say for theories of objective permissibility.

Ultimately, neither of the standards of subjective permissibility considered in this paper is *really* suited to decision-making under pressure, when lives are at stake. The Threshold View looked simple, but it offers only a necessary condition for subjective permissibility; working out what one is subjectively permitted to do will be much more complex, especially once proportionality and necessity come into the

picture. The Decision-Theoretic Standard too would be hard to apply. But perhaps moral philosophers cannot hope for more. Indeed, it seems unlikely that there is any single decision procedure that will plausibly apply to all individuals, in all circumstances. What works for one stoical, cool-headed natural mathematician might be hopeless for a quick-tempered person with dyscalculia. In seeking a standard of subjective permissibility, we cannot hope to give people a foolproof way to make moral decisions.

What we can aim to do, however, is to pick out the option that a morally skilled agent would choose in the circumstances. When we are uncertain about crucial facts, moral skill does not consist simply in doing what is objectively right, because to do the objectively right thing when the facts are in doubt involves luck, not skill. The morally skilled agent would adhere to the Decision-Theoretic Standard because it represents the best way of balancing risks and opportunities when deciding under uncertainty. Morally skilled agents will not always do the expected moral utility calculations—more commonly, they will make a judgement call, perceiving the relevant trade-offs intuitively. But when we later review their decisions, we will find that the option chosen was the one that would have been selected by applying this standard. Importantly, this also gives us a measure by which to calibrate the more practical decision procedures that we do come up with. They succeed insofar as they enable realistic agents to pick out the option also picked out by the Decision-Theoretic Standard.

One might still ask: why would the morally skilled agent maximise expected moral utility? Why does this constitute the best trade-off between risks and opportunities? On my naïve, or classical approach, I cannot support these judgements with representation theorems and their underpinning axioms. I can only make a simple observation. If objective permissibility has to do with the reasons that apply to our actions, and if those reasons can vary depending on what state the world happens to be in, then it would be absurd to treat every possible reason as though it were actual. Our criterion of subjective permissibility must therefore discount reasons for the probability that the world is in that state. The requirement that we maximise expected moral utility, subject to the reasonable costs proviso, simply enjoins us to discount those reasons in proportion to their probability of being actual. What other discount makes more sense? The burden of proof is on the opponent to explain just why some other discount, out of proportion to probability, would be superior.[34]

5 Conclusion

Perhaps moral philosophers can tell those who deal in life and death very little. Certainly, given how radically most of us are removed from such exigent situations, we should wait until we are asked. Then, we can say more than simply 'do what is

---

[34] One possible avenue: argue that morality includes some in-built risk-aversion (or risk-seeking) like that modelled in Lara Buchak's (2013) risk-weighted expected utility theory.

objectively right'. We can say: of the options that have reasonable expected costs, you may choose the one that maximises expected moral utility. The tools of decision theory are available to us all, deontologists included, once we understand its different components favourably.

This is only a starting point: the real challenge for this approach comes in specifying just how to model one's moral theory for the purposes of decision-making under uncertainty. But even before we do so, this theoretical conclusion has real practical import: those who must use lethal force under uncertainty cannot ignore the stakes, and focus only on probability of liability, as the Threshold View would have them do. When the stakes are high enough, they can be subjectively permitted to use lethal force, even if they are not sufficiently confident of their target's liability to be killed. Rules like those espoused by the Obama administration might be a sensible heuristic, but they do not describe an underlying moral truth.

Of course, there is no single decision theory, but rather a number of decision theories, and exploring the intricacies of the field will lead to many refinements of this central principle. Indeed, the next step in this project is to identify elements in deontological ethics that resist straightforward incorporation within the Decision-Theoretic Standard, and force us to introduce substantial modifications—for example elsewhere I identify agent-centred options and anti-aggregationist principles, as posing significant challenges (Lazar 2017a, b). And further refinements might come from within decision theory itself.[35] There is a rich seam to be mined by ethicists and decision theorists working together. In this paper, I have sought only to take the first steps—to ward deontologists away from the threshold approach, and to show that endorsing a deontological approach to the ethics of killing does not preclude offering a plausible account of what one may do when lives are at stake and the facts are in doubt.

References

Aboodi, R., Borer, A., & Enoch, D. (2008). Deontology, individualism, and uncertainty. *Journal of Philosophy, 105*(5), 259–272.
Bazargan, S. (2014). Killing minimally responsible threats. *Ethics, 125*(1), 114–136.
Broome, J. (1991). *Weighing goods: Equality, uncertainty and time*. Oxford: Blackwell.

---

[35] See, for example, (Nover and Hájek 2004).

Buchak, L. (2013). *Risk and rationality*. Oxford: Oxford University Press.

Buchak, L. (2016). Decision Theory. In A. Hájek & C. Hitchcock (Eds.), *Oxford handbook of the philosophy of probability*. Oxford: Oxford University Press.

Colyvan, M., Cox, D., & Steele, K. (2010). Modelling the moral dimension of decisions. *Noûs, 44*(3), 503–529. doi:10.1111/j.1468-0068.2010.00754.x.

Cook, R. T. (2002). Vagueness and mathematical precision. *Mind, 111*(442), 225–247.

Dougherty, T. (2013a). Agent-neutral deontology. *Philosophical Studies, 163*(2), 527–537.

Dougherty, T. (2013b). Aggregation, beneficence and chance. *Journal of Ethics and Social Philosophy, 7*(2), 1–19.

Dougherty, T. (2013c). Vague value. *Philosophy and Phenomenological Research, 89*(2), 352–372. doi:10.1111/phpr.12026.

Dreier, J. (1993). Structures of normative theories. *The Monist, 76*(1), 22–40.

Dreier, J. (2004). Why ethical satisficing makes sense and rational satisficing doesn't. In M. Byron (Ed.), *Satisficing and maximizing: Moral theorists on practical reason* (pp. 131–154). Cambridge: Cambridge University Press.

Dreier, J. (2011). In defense of consequentializing. In M. Timmons (Ed.), *Oxford studies in normative ethics* (Vol. 1, pp. 97–118). Oxford: Oxford University Press.

Easwaran, K. (2014). Decision theory without representation theorems. *Philosophers' Imprint, 14*(27), 1–30.

Feldman, F. (2006). Actual utility, the objection from impracticality, and the move to expected utility. *Philosophical Studies, 129*(1), 49–79. doi:10.1007/s11098-005-3021-y.

Ferzan, K. (2005). Justifying self-defense. *Law and Philosophy, 24*(6), 711–749.

Firth, J. M., & Quong, J. (2012). Necessity, moral liability, and defensive harm. *Law and Philosophy, 31*(6), 673–701. doi:10.1007/s10982-012-9135-4.

Fried, B. H. (2012). What does matter? The case for killing the trolley problem (or Letting It Die). *Philosophical Quarterly, 62*(248), 505–529.

Frowe, H. (2010). A practical account of self-defence. *Law and Philosophy, 29,* 245–272.

Frowe, H. (2014). *Defensive killing*. Oxford: Oxford University Press.

Graham, P. A. (2010). In defense of objectivism about moral obligation. *Ethics, 121*(1), 88–115.

Haque, A. A. (2012). Killing in the fog of war. *Southern California Law Review, 86*(1), 63–116.

Hawley, P. (2008). Moral absolutism defended. *Journal of Philosophy, 105*(5), 273–275.

Hedden, B. (2016). Does mite make right? On decision-making under normative uncertainty. *Oxford Studies in Metaethics, 11,* 102–128.

Huemer, M. (2010). Lexical priority and the problem of risk. *Pacific Philosophical Quarterly, 91*(3), 332–351.

Hurley, P. (2013). *Consequentializing and deontologizing: Clogging the consequentialist vacuum*. Oxford: Oxford University Press.

Isaacs, Y. (2014). Duty and knowledge. *Philosophical Perspectives, 28*(1), 95–110. doi:10.1111/phpe.12042.

Jackson, F. (1991). Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics, 101*(3), 461–482.

Jackson, F., & Smith, M. (2006). Absolutist moral theories and uncertainty. *Journal of Philosophy, 103*(6), 267–283.

Jackson, F., & Smith, M. (2015). The implementation problem for deontology. In B. Maguire & E. Lord (Eds.), *Weighing reasons* (pp. 279–291). Oxford: Oxford University Press.

Jeffrey, R. C. (1965). *The logic of decision (Mcgraw-Hill Series in Probability and Statistics)*. New York: McGraw-Hill.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty*. Cambridge: Cambridge University Press.

Kamm, F. M. (2007). *Intricate ethics: Rights, responsibilities, and permissible harm (Oxford Ethics Series)*. New York: Oxford University Press.

Lazar, S. (2012). Necessity in self-defense and war. *Philosophy & Public Affairs, 40*(1), 3–44.

Lazar, S. (2013). Associative duties and the ethics of killing in war. *Journal of Practical Ethics, 1*(1), 3–48.

Lazar, S. (2015a). Authority, oaths, contracts, and uncertainty in war. *Thought: A Journal of Philosophy, 4*(1), 52–58. doi:10.1002/tht3.157.

Lazar, S. (2015b). Risky killing and the ethics of war. *Ethics, 126*(1), 91–117.

Lazar, S. (2016). Anton's game: Deontological decision theory for an iterated decision problem. *Utilitas*. doi:10.1017/S0953820816000236.

Lazar, S. (2017a). Aggregation and risk. *Unpublished MS*.

Lazar, S. (2017b). Deontological decision theory and agent-centred options. *Ethics, 127*(3), 579–609.

Lazar, S. (2017c). Risky killing: How risks worsen objective rights-violations. *Journal of Moral Philosophy*. doi:10.1163/17455243-46810076.

Lazar, S., & Lee-Stronach, C. (2016). Axiological absolutism and risk. *Unpublished MS*.

Lin, H. (2013). Foundations of everyday practical reasoning. *Journal of Philosophical Logic, 42*(6), 831–862.

Louise, J. (2004). Relativity of value and the consequentialist umbrella. *The Philosophical Quarterly, 54*(217), 518–536.

McCarthy, D. (1996). Liability and risk. *Philosophy & Public Affairs, 25*(3), 238–262.

McCarthy, D. (1997). Rights, explanation, and risks. *Ethics, 107*(2), 205–225.

McMahan, J. (1994). Innocence, self-defense and killing in war. *Journal of Political Philosophy, 2*(3), 193–221. doi:10.1111/j.1467-9760.1994.tb00021.x.

McMahan, J. (2009). *Killing in war (Uehiro Series in Practical Ethics)*. Oxford: Oxford University Press.

McMahan, J. (2011). Who is morally liable to be killed in war? *Analysis, 71*(3), 544–559. doi:10.1093/analys/anr072.

McMahan, J. (2014). What rights may be defended by means of war? In S. Lazar & C. Fabre (Eds.), *The morality of defensive war* (pp. 115–158). Oxford: Oxford University Press.

McNaughton, D., & Rawling, P. (1991). Agent-relativity and the doing-happening distinction. *Philosophical Studies, 63*(2), 167–185. doi:10.2307/4320228.

McNaughton, D., & Rawling, P. (1998). On defending deontology. *Ratio, 11*(1), 37–54.

Meacham, C. J. G., & Weisberg, J. (2010). Representation theorems and the foundations of decision theory. *Australasian Journal of Philosophy, 89*(4), 641–663. doi:10.1080/00048402.2010.510529.

Norcross, A. (1997). Comparing harms: Headaches and human lives. *Philosophy & Public Affairs, 26*(2), 135–167. doi:10.1111/j.1088-4963.1997.tb00079.x.

Nover, H., & Hájek, A. (2004). Vexing expectations. *Mind, 113*(450), 237–249.

Oddie, G., & Milne, P. (1991). Act and value: Expectation and the representability of moral theories. *Theoria, 57*(1–2), 42–76.

Office, W. H. P. (2013). Fact sheet: U.S. policy standards and procedures for the use of force in counterterrorism operations outside the United States and areas of active hostilities.

Parfit, D. (2011). *On what matters*. Oxford: Oxford University Press.

Peterson, M. (2002). An argument for the principle of maximizing expected utility. *Theoria, 68*(2), 112–128.

Peterson, M. (2010). A royal road to consequentialism? *Ethical Theory and Moral Practice, 13*(2), 153–169.

Portmore, D. W. (2009). Consequentializing. *Philosophy Compass, 4*(2), 329–347. doi:10.1111/j.1747-9991.2009.00198.x.

Portmore, D. W. (2016). Uncertainty, indeterminacy, and agent-centred constraints. *Australasian Journal of Philosophy*. doi:10.1080/00048402.2016.1219376.

Quong, J. (2015). Rights against harm. *Aristotelian Society Supplementary, 89*(1), 249–266. doi:10.1111/j.1467-8349.2015.00252.x.

Rawls, J. (1999). *A theory of justice* (Rev ed.). Oxford: Oxford University Press.

Raz, J. (2011). *From normativity to responsibility*. Oxford: Oxford University Press.

Regan, D. (1982). *Utilitarianism and cooperation*. Oxford: Oxford University Press.

Ross, J. (2006). Rejecting ethical deflationism. *Ethics, 116*(4), 742–768.

Savage, L. J. (1954). *The foundations of statistics (Wiley Publications in Statistics)*. New York: Wiley.

Scheffler, S. (1994). *The rejection of consequentialism*. Oxford: Clarendon Press.

Schroeder, M. (2007). Teleology, agent-relative value, and "Good". *Ethics, 117*(2), 265–295. doi:10.1086/511662.

Schroeder, S. A. (2011). You don't have to do what's best! A problem for consequentialists and other teleologists. In M. Timmons (Ed.), *Oxford studies in normative ethics* (Vol. 1, pp. 166–201). Oxford: Oxford University Press.

Schroeder, S. A. (Forthcoming). Consequentializing and its consequences. *Philosophical Studies*, 1–23.

Smith, M. (2007). Meta-Ethics. In F. Jackson & M. Smith (Eds.), *The oxford handbook of contemporary philosophy* (pp. 3–30). Oxford: Oxford University Press.

Smith, M. (2009). Two kinds of consequentialism. *Philosophical Issues, 19*(1), 257–272. doi:10.1111/j.1533-6077.2009.00169.x.

Spector, H. (2016). Decisional nonconsequentialism and the risk sensitivity of obligation. *Social Philosophy and Policy, 32*(2), 91–128.

Tenenbaum, S. (2014). The perils of earnest consequentializing. *Philosophy and Phenomenological Research, 88*(1), 233–240. doi:10.1111/phpr.12080.

Vallentyne, P. (1988). Gimmicky representations of moral theories. *Metaphilosophy, 19*(3–4), 253–263. doi:10.1111/j.1467-9973.1988.tb00772.x.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Walen, A. (2015). Proof beyond a reasonable doubt: A balanced retributive account. *Louisiana Law Review, 76*(2), 355–446.

Weirich, P. (2004). *Realistic decision theory: Rules for nonideal agents in nonideal circumstances*. New York: Oxford University Press.

Zimmerman, M. J. (2008). *Living with uncertainty: The moral significance of ignorance*. Cambridge: Cambridge University Press.