Morality Under Risk

Chad Lee-Stronach

A thesis to be submitted for the degree of
Doctor of Philosophy of
The Australian National University

October 2018

Abstract

Many argue that absolutist moral theories – those that prohibit particular kinds of actions or trade-offs under all circumstances – cannot adequately account for the permissibility of risky actions. In this dissertation, I defend various versions of absolutism against this critique, using overlooked resources from formal decision theory. Against the prevailing view, I argue that almost all absolutist moral theories can give systematic and plausible verdicts about what to do in risky cases. In doing so, I show that critics have overlooked: (1) the fact that absolutist theories – and moral theories, more generally – underdetermine their formal decision-theoretic representations; (2) that decision theories themselves can be generalised to better accommodate distinctively absolutist commitments. Overall, this dissertation demonstrates that we can navigate a risky world without compromising our moral commitments.

Contents

Introduction

0.1 Morality in a Risky World

The world is a risky place. Everyday actions like driving, making a promise, or purchasing a product can – for all we know – cause harm, be insincere, or sustain others' wrongful practices. In light of this fact, it is surprising that few moral theories provide us any systematic guidance about which kinds of risky actions are morally permissible, and which are not.[1] This is a problem: given that we are never completely certain about the nature or consequences of our actions, it seems that our moral theories are, for all practical purposes, silent.[2]

Recognising this problem, various philosophers have attempted to extend our existing moral theories to give verdicts in cases involving uncertainty. However, they have found that some kinds of moral theories – so-called *absolutist* moral theories – give unacceptable verdicts in such cases, and should therefore be rejected. I shall refer to this critique as:

The Problem of Risk: Absolutist moral theories cannot give adequate moral guidance in cases of uncertainty.

In this dissertation, I argue that the Problem of Risk is false. Using overlooked resources from formal decision theory, I show how various absolutist theories can provide adequate guidance through an risky world.

The overall structure of the dissertation is straightforward. **Part 1** introduces the debate over Moral Absolutism and risk. To make the debate more tractable, I define a set of necessary and sufficient conditions for a theory to give adequate moral guidance in cases of uncertainty. I argue that most absolutist theories can satisfy these conditions. **Part 2** details ways they might do so.

1. Here and throughout, I use the terms 'uncertain' and 'risky' interchangeably, assuming a broadly Bayesian approach that allows for assignments of probabilities (be precise or imprecise, objective or subjective). On these issues, see: Alan Hájek and Christopher Hitchcock, eds., *The Oxford Handbook of Probability and Philosophy* (Oxford: Oxford University Press, 2016).

2. Here and throughout, I will focus solely on empirical uncertainty: uncertainty about an action's kind or consequences, holding fixed a background moral theory. This is in contrast with moral uncertainty, which concerns uncertainty about which moral theory is true. See: Andrew Sepielli, "What to Do When You Don't Know What to Do," in *Oxford Studies in Metaethics* (2009).

0.2 What is 'Moral Absolutism'?

'Moral Absolutism' is something of a pejorative label. It is usually applied to moral theories on the basis of some perceived fanaticism or dogmatism. For instance, Michael Huemer (2010) takes the following theorists to be absolutists:

> Kant held that one must always keep one's promises, no matter how bad the consequences of doing so may be, or how much good might be brought about by breaking a promise. Elizabeth Anscombe held that it is always wrong to knowingly punish a person for a crime he did not commit. And Robert Nozick appears to have held that it is always wrong to violate an individual's negative rights against coercion. [3]

More formally, Frank Jackson and Michael Smith (2006) hold that Moral Absolutism is any theory that:

> Absolutely prohibits actions of kind K, where K ... is a property of an action as opposed to a relation between an action and available alternatives to that action:[4]

They, like others, have since expanded the category of Absolutism to include deontology (both moderate and strict forms), rights-theories, lexical priority theories, anti-aggregationist theories, and those that take some actions to be unconditionally wrongful. These and other 'absolutist' theories are said to face the Problem of Risk.

For instance, Robert Nozick (1974) discusses the Problem of Risk as it relates to natural rights theories.[5] Dennis McKerlie (1986) reinforces Nozick's worries, coming to the conclusion that rights-based theories cannot operate in cases of risk.[6] J.E.J. Altham (1983) extends the Problem to other rights-based theories, including contractualism.[7] Judith Jarvis Thomson (1986) argues that theories that posit a right against risk lead to unacceptable verdicts about whether an action is rights-violating.[8] Shelly Kagan (1991) explores how deontological constraints – roughly, limitations on how we can make the world a better place – operate in risky cases. He is unsatisfied with any approach to dealing with risk, along similar lines to those given by McKerlie, arguing that either deontology takes all, none, or some

3. Michael Huemer, "Lexical Priority and the Problem of Risk," *Pacific Philosophical Quarterly* 91 (2010): p. 332.

4. Frank Jackson and Michael Smith, "Absolutist Moral Theories and Uncertainty," *The Journal of Philosophy* 103, no. 6 (2006): p. 267.

5. Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic, 1974).

6. Dennis McKerlie, "Rights and Risk," *Canadian Journal of Philosophy* 16, no. 2 (1986): 239–251.

7. J. E. J. Altham, "Ethics of Risk," *Proceedings of the Aristotelian Society, New Series* 83 (1983): 15–29.

8. Judith Jarvis Thomson, *Rights, Restitution, and Risk: Essays in Moral Theory* (Cambridge, Mass.: Harvard University Press, 1986), p. 177, 181.

subset (defined by some arbitrary threshold) of risks to be wrongful.[9] More recently, this trilemmatic approach was finessed by Frank Jackson and Michael Smith (2006).[10] Similarly, Sven Ove Hansson (2003, 2013) argues that almost all contemporary moral theories face the Problem of Risk (which he calls the "Mixture Appraisal Problem").[11] Michael Huemer (2010) argues that the Problem of Risk applies to all lexical priority theories: those that prioritise some moral considerations over any number of particular other considerations.[12] Jackson and Smith (2016) argue that all versions of deontology face the Problem.[13] Yoaav Isaacs (2014) agrees and proposes, but ultimately does not endorse, a knowledge-first approach for deontology to avoid the Problem of Risk.[14] Holly Smith (2014) develops a specific version of the Problem of Risk, arguing that deontologists cannot, by their own lights, be required to gather more information before acting.[15] Collectively, these critiques lead us to the conclusion that a vast swathe of moral theorising cannot guide us through an uncertain world, and should therefore be rejected.

Why do these theories face the Problem of Risk? I shall argue that the root cause is structural, not substantive. By substantive, I mean *what* a moral theory considers to be important: life, love, liberty, what have you. By structural, I mean the logical relations that connect these substantive considerations. As we shall see, the Problem of Risk relies on attributing to a moral theory particular structural assumptions: specifically, assumptions that prevent the theory's substantive considerations from being sufficiently measurable or comparable. Over the course of this dissertation, I shall argue in detail that many allegedly absolutist theories are not committed to the structural assumptions that give rise to the Problem of Risk. In general, I will not enter a dispute about labels: from the vantage point of some other moral theory, each of these could well seem absolutist in some relevant sense. However, I will show that even if these moral theories are 'absolutist', that does not commit them to the Problem of Risk. In fact, I will attempt to show that practically all of the allegedly absolutist theories mentioned above can operate satisfactorily in a risky world.

9. Shelly Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1991), pp. 87-92.

10. Jackson and Smith, "Absolutist Moral Theories."

11. Sven Ove Hansson, "Ethical Criteria of Risk Acceptance," *Erkenntnis* 59, no. 3 (2003): 291–309; Sven Ove Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World* (Palgrave Macmillan, 2013).

12. Huemer, "Lexical Priority and the Problem of Risk."

13. Errol Lord and Barry Maguire, eds., "The Implementation Problem for Deontology," chap. 14 in *Weighing Reasons* (Oxford: Oxford University Press, 2016), 338–354.

14. Yoaav Isaacs, "Duty and Knowledge," *Philosophical Perspectives* 28, no. 1 (2014): 95–110.

15. Holly M. Smith, "The Subjective Moral Duty to Inform Oneself before Acting," *Ethics* 125, no. 1 (2014): 11–38.

0.3 Methodology

How can we determine whether absolutist theories can avoid the Problem of Risk? I believe that the history of cartography provides an answer.[16] For the longest time, maps were highly unreliable. This is because explorers often used intuition to mark the landmarks and the distances between them. Intuition is, of course, a highly fallible guide to such matters, leading to mistaken judgements that differed from person to person, and allowing for *ad hoc* details to creep in to fill in the gaps left by intuition or observation. Jonathan Swift skewered such explorers:

> So Geographers, in Afric-maps,
> With savage-pictures fill their gaps;
> And o'er uninhabitable downs
> Place elephants for want of towns.[17]

The great advances in cartography came with the use of trigonometric methods and observational and measuring techniques. Using a set of axioms defining the relationships between distance and angles, alongside more advanced methods of measuring each, cartographers were reliably guided beyond their intuitions, providing a consistent and increasingly accurate representation of largely untouched terrain.

Today, moral philosophy still operates on the basis of intuitions about what is right and wrong under hypothetical, 'unexplored' scenarios. For the most part, however, this is a reliable approach. Unfortunately, in cases of uncertainty, this methodology is fraught: psychologists have persuasively shown that our intuitive judgements about risk are often naive or conflicting, or otherwise subject to one of many cognitive biases.[18] To advance moral theorising over these difficult cases, we must find an equivalent to the trigonometric methods used by cartographers, one that allows us to project our understanding from the firmer grounds of certainty to the cognitively distant terrains of uncertainty.

Not only do we need a method that does not rely heavily on intuition, we also need one that allows for a sufficient degree of measurability and comparability. Without a method for systematically distinguishing the *degrees* to which risky actions are right or wrong, moral philosophy will run afoul of a basic cartographic principle, noted in the third century AD by Pei Xiu:

> If one draws a map without having graduated divisions, there is no means of distinguishing between what is near and what is far.[19]

16. John Noble Wilford, *The Mapmakers*, Revised Edition (New York: Vintage Books, 2000).
17. ibid., p. 14.
18. We tend to overweight small probabilities, make mistakes when conjoining probabilities, mistake salience with probability, among many other errors.Amos Tversky and Daniel Kahneman, "Judgement Under Uncertainty: Heuristics and Biases," *Science* 185, no. 4157 (1974): 1124–1131; Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus / Giroux, 2011).
19. Wilford, *The Mapmakers*, p. 33.

I shall argue that formal theories of rational decision give us a useful way of systematically projecting our moral commitments beyond the *terra firma* of our strongly-held intuitions and into the domain of uncertainty, providing a more systematic and reliable way of charting the moral dimensions of risky cases. Specifically, I will argue that most absolutists can avoid the Problem of Risk by adopting some version of expected value theory. Expected value theory defines sets of axioms governing how we ought to treat the relationships between value, probability, and actions. Specifically, it holds that we should choose actions that maximise probability-weighted average value. In general, I will adopt broadly von Neumann-Morgenstern axioms, while also allowing for violations of its axioms of Continuity (allowing for a prohibition of particular kinds of trade-offs) and Completeness (allowing for the incommensurability of particular moral considerations).[20] At times, I will adopt a formally more general decision-theoretic approach, based on Dietrich and List (2017).[21] These formal frameworks are by no means the only ways of modelling absolutism – they simply suffice to frame the discussions and to provide positive proposals for how absolutists can avoid the Problem of Risk.

Why take this approach as opposed to some informal criterion, such as 'reasonableness' or 'foreseeability'? Firstly, expected value theory can be more *informative* in explaining why one risky action is permissible, yet another is not: this will be due to the values or probabilities (or some combination thereof) of the various options available. Informal standards, by contrast, often obscure these variables or else 'pass the buck' by lapsing into disguised tautology: e.g. risks are permissible when they are 'reasonable' or 'acceptable'.[22]

Secondly, expected value theory is *comprehensive* in that provides verdicts for all combinations of probability and values.[23] Provided that we accept the axioms and value assignments, we can be confident in its verdicts across a range of cases which

20. John Von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior*, 3rd (1947; Princeton: Princeton University Press, 1990). Since these axioms are defined in terms of objective probabilities, I will implicitly supplement them with subjectivised axioms, as per: F. J. Anscombe and R. J. Aumann, "A Definition of Subjective Probability," *The Annals of Mathematical Statistics* 34, no. 1 (1963): 199–205. Alternative approaches that could have been used include: Leonard J. Savage, *The Foundations of Statistics*, 2nd rev. (New York: Dover, 1972); Richard C. Jeffrey, *The Logic of Decision*, Second (Chicago and London: The University of Chicago Press, 1983).

21. Franz Dietrich and Christian List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories," *The Philosophical Review* 126, no. 4 (2017): 421–479.

22. For example, Scanlon (2008) appeals to 'reasonable' belief to justify harm done in self-defence, but does not explain what a reasonable belief is. Hansson (2003) explictly rejects expected value approaches to risk, instead taking 'acceptable' risks to be those that are part of a 'persistently justice-seeking' social system. Unfortunately, he does not venture to say what it means for a risk to be 'part of' such a system. T.M. Scanlon, *Moral Dimensions* (Cambridge Massachusetts, and London, England: Belknap Press of Harvard University Press, 2008), p. 67; Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World*, Ch. 6.

23. Okay, maybe not *all* possible combinations: Harris Nover and Alan Hájek, "Vexing Expectations," *Mind* 113, no. 450 (2004): 237–249.

we haven't specifically considered – or indeed could not consider (with any degree of accuracy). Expected value theory is thus *ampliative* in the sense of providing verdicts beyond those already reached by existing theorising. Informal standards, by contrast, are rarely ampliative – we cannot rely on their verdicts about risky cases without having to directly rely on our questionable intuitions about them.

Thirdly, expected value theory is provably *consistent*: properly applied, it does not lead to a contradictory set of verdicts.[24] By contrast, those who avoid decision-theoretic approaches in favour of piecemeal, case-by-case principles, cannot be certain that these principles will yield a consistent set.[25]

Lastly, adopting this kind of formal approach allows us to better understand the commitments of moral theories and the sources of disputes between them. If there is disagreement about the acceptability of a model's verdicts, then the dispute lies in the assignment of values or in one or more of the axioms. In this way, the decision-theoretic approach helps to isolate the points of dispute. As we shall see, it also helps to reveal previously overlooked aspects of the moral theories we will consider.[26]

Based on these methodological considerations, I will adopt an approach known as *decision-theoretic ethics*.[27] This method involves first determining whether a given moral theory is logically compatible with some version of decision theory. If it is, then the challenge is to demonstrate that modelling the moral theory using decision theory does not lead it to give unacceptable verdicts about the permissibility of risky actions.[28] From the point of view of this method, the Problem of Risk can be understood as the claim that some types of moral theories are either incompatible with decision theory or else give palpably incorrect results when they are extended, via decision theory, to cover risky cases.

Before continuing, however, I should note that others have come to the defence of these various moral theories, many using broadly decision-theoretic approaches. For example, following arguments that rights-based theories have difficulty accommodating risk, David McCarthy (1997) argues that such theories are compatible with orthodox decision theory.[29] However, it seems that this response has been largely

24. This has been proven by a range of representation theorems, which I need not discuss here. See: Von Neumann and Morgenstern, *Theory of Games and Economic Behavior*; Savage, *The Foundations of Statistics*; Jeffrey, *The Logic of Decision*.

25. On this point, see: Chapter 1, Section 3, Subsection 1; Chapter 4, Section 4.

26. Thanks here to Christian List.

27. Frank Jackson, "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection," *Ethics* 101, no. 3 (1991): 461–482.

28. Some moral theorists appear to believe that it is open to them to reject decision theory. However, I will attempt to show that this involves rejecting very minimal and compelling axioms of moral decision-making. More generally, I hope to show that the conflict between some kinds of moral theories and expected value theory has been exaggerated, such that moral theorists do not need to give up expected value theory at all.

29. David McCarthy, "Rights, Explanation, and Risks," *Ethics* 107, no. 2 (1997): 205–225.

overlooked.[30]

Deontologists, more generally, have recently begun exploring these limits of expected value theory in modelling their distinctive moral commitments, such as agent-centred perogatives and options.[31] Some have argued that deontologists should reject expected value theory entirely.[32] I will not explore whether existing decision-theoretic models can accommodate agent-centred perogatives and options, but I will suggest that at least some deontological theories can benefit from the generalised decision-theoretic approaches I canvass here.

Others have attempted to model broadly absolutist theories without rejecting any axioms of expected value theory.[33] However, such approaches often require complex interpretations of the formal framework to prevent violations of the axioms.[34] For instance, deontological theories that try to hew too closely to orthodox decision theory (usually, by some acrobatic book-keeping) arguably trivialise the model, making it difficult to give principled determinations of whether individuals ever violate deontological morality.[35]

As we shall see, my approach often differs from the above by taking a middle-ground between all-out rejecting or accepting orthodox decision theory. I often adopt generalised models that allow for violations of some of the standard axioms of expected value theory or Bayesian epistemology. My contributions can thus be read as supplementing, rather than supplanting, those made elsewhere. In general, I will adopt a methodologically conservative approach of attempting to provide a decision-theoretic formal model that minimally departs from orthodox expected value theory, without trivialising the model. This approach optimises the explanatory power and systematicity of decision theory, providing moral theories an ampliative structure,

30. For instance, see the literature on the Paralysis Problem: Madeleine Hayenhjelm and Jonathan Wolff, "The Moral Problem of Risk Impositions: A Survey of the Literature," *European Journal of Philosophy* 20, no. 51 (2011): E26–E51; Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World*. For what seems to have been the only published response to McCarthy's proposal, see: Stephen Perry, "Risk, Harm, Interests, and Rights" [in eng], in *Risk: Philosophical Perspectives*, ed. Tim Lewens (Routledge: Oxford University Press, 2007).

31. Seth Lazar, "Deontological Decision Theory and Agent-Centered Options," *Ethics* 127, no. 3 (2017): 579–609; Seth Lazar, "In Dubious Battle: Uncertainty and the Ethics of Killing," *Philosophical Studies*, 2017, 1–27; Seth Lazar, "Anton's Game: Deontological Decision Theory for an Iterated Decision," *Utilitas* 29, no. 1 (2017): 88–109; Seth Lazar, "Risky Killing: How Risks Worsen Violations of Objective Rights," *Journal of Moral Philosophy*, 2017, 1–30; Seth Lazar, "Limited Aggregation and Risk" (2018), 1–28; Douglas W Portmore, *Commonsense Consequentialism: Wherein Morality Meets Rationality* (Oxford: Oxford University Press, 2011), xx, 266 p.

32. Sergio Tenenbaum, "Action, Deontology, and Risk: Against the Multiplicative Model," *Ethics* 127 (2017): 1–36; Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World*.

33. Graham Oddie and Peter Milne, "Act and Value: Expectation and the Representability of Moral Theories," *Theoria*, 1991, Mark Colyvan, Damian Cox, and Katie Steele, "Modelling the Moral Dimension of Decisions," *Noûs* 44, no. 3 (September 2010): 503–529

34. Larry S. Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Oxford; New York: Oxford University Press, 2012), Ch. 13.

35. John Broome, *Weighing Goods* (Cambridge, Massachusetts: Blackwell, 1995), pp. 103-106; Lazar, "Deontological Decision Theory and Agent-Centered Options."

without compromising their underlying substantive commitments.

0.4 Outline of the Dissertation

This project is broadly a positive one in the sense that it *negates* negative results that have been pressed against absolutist moral theories. This involves presenting a series of 'proofs of possibility' in response to particular discussions where some or other theory is believed to be subject to the Problem of Risk. I take these results to show that accommodating risk in our moral theories does not require us to make any important substantive concessions about what matters, morally-speaking, or why. The dissertation has two main parts:

0.4.1 Part 1: The Problem of Risk

In Chapter 1: Prohibition and Probability, I give a formal representation of the structural features of a moral theory that make it vulnerable to the Problem of Risk. I establish that the Problem only applies to a very particular version of Moral Absolutism, namely: Option Absolutism, which holds that some actions are prohibited irrespective of their alternatives. I then argue that few, if any, existing absolutist theories are of this kind. Instead, such theories are better understood as subscribing to Relational Absolutism, which holds that an action's being prohibited depends on its relation to other available actions. I show that Relational Absolutist theories can potentially avoid the Problem of Risk. This clears the way for positive proposals in later chapters, which detail how various other absolutist theories might avoid the Problem of Risk.

0.4.2 Part 2: Solutions

In light of the possibility result given in Chapter 1, I set out to develop various 'proofs of possibility' that show how particular moral theories can potentially operate in risky cases, despite their being 'absolutist' in some broad sense.

Chapter 2: Moral Priorities Under Risk, defends lexical priority theories using lexicographic expected value theory.[36] This chapter illustrates that lexical priority theories can be understood as being Relational Absolutist. It also identifies and responds to various challenges to lexical priority theories.

Chapter 3: Priorities and Uncertainties (closely based on a co-authored paper with Seth Lazar), provides an account of how standard expected value theory can

36. This chapter is very close to its published version: Chad Lee-Stronach, "Moral Priorities Under Risk," *Canadian Journal of Philosophy*, 2017, 1–19

allow lexical priority theories to avoid the Problem of Risk.[37] This chapter also reframes the challenge as not concerning risk as such, but rather other substantive issues such as whether individual or sequences of actions are the proper objects of moral evaluation.

Chapter 4: Duty and Ignorance considers a different version of the Problem of Risk, given by Holly Smith (2014), which I call the Problem of Ignorance.[38] It concerns cases where a moral theory makes incorrect verdicts about whether you ought to gather more information before acting. I respond to Smith's arguments that deontological theories face the Problem of Ignorance. I present a decision-theoretic model of deontological theories that avoids the Problem.

Lastly, Chapter 5: Authority, Obedience and Uncertainty, considers the Problem of Risk in relation to the literature on legitimate authority: specifically, Joseph Raz's Service Conception of authority.[39] Raz's theory is relevant because it appears to exhort an absolutist-like obedience to legitimate authority. I argue that Raz's theory can be satisfactorily extended to cases of uncertainty using expected value theory. I then respond to objections to decision-theoretic approaches to legitimate authority.[40]

These chapters are designed to be separate interventions into particular debates over the Problem of Risk. Taken together, they illustrate that the Problem of Risk is a productive challenge to moral theories, but one that is not likely to be successful. Once we clarify a moral theory's substantive commitments and take a full view of the decision-theoretic resources available, we can see that most – if not all – 'absolutist' theories can guide us through a risky world.

Having charted the course forward into the terrain of uncertainty, we must now explore these relatively untouched grounds up close.

Onwards!

37. Apart from terminological changes and slight changes of focus, this chapter closely follows its published version: Seth Lazar and Chad Lee-Stronach, "Axiological Absolutism and Risk," *Noûs*, 2017, 1–17.

38. Smith, "The Subjective Moral Duty to Inform Oneself before Acting."

39. Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986).

40. Scott J. Shapiro, *Authority*, ed. Jules L. Coleman, Kenneth Einar Himma, and Scott J. Shapiro (Oxford University Press, 2004), 382–439.

Part I

The Problem of Risk

Chapter 1

Prohibition and Probability

1.1 Introduction

Most contemporary moral theories hold that the nature or consequences of an action can render it morally permissible or impermissible.[1] Absolutist moral theories hold that some actions are *always* morally impermissible. For example, the following theorists are taken to be prototypical moral absolutists:

> Kant held that one must always keep one's promises, no matter how bad the consequences of doing so may be, or how much consideration might be brought about by breaking a promise. Elizabeth Anscombe held that it is always wrong to knowingly punish a person for a crime he did not commit. And Robert Nozick appears to have held that it is always wrong to violate an individual's negative rights against coercion.[2]

In recent years, many have argued that absolutist theories have very little to say about what we ought to do when we are, to some degree, *uncertain* about whether a given action is of the absolutely prohibited kind.[3] Moreover, they have argued that when such theories are extended to cover such cases, they seem to give incorrect verdicts. I shall call this critique:

The Problem of Risk: Absolutist moral theories cannot give adequate moral guidance in cases involving risk.

The Problem of Risk has been presented as a challenge for various, broadly absolutist, moral theories.[4] However, it has never been precisely explained why some

1. For a recent theory that avoids such moral permissibility verdicts, see: Caspar Hare, *The Limits of Kindness* (Oxford: Oxford University Press, 2013).

2. Huemer, "Lexical Priority and the Problem of Risk," p. 332.

3. For example, see: Jackson and Smith, "Absolutist Moral Theories."

4. For instance, the Problem of Risk led Judith Jarvis Thomson (1986, 1990) and Dennis McKerlie (1986) to conclude that we do not have rights against having risk imposed on us. See: Thomson, *Rights, Restitution, and Risk: Essays in Moral Theory*; Judith Jarvis Thomson, *The Realm of*

moral theories seem to encounter the Problem, while others seem to avoid it. This paper identifies a set of structural assumptions that give rise to the Problem of Risk. It argues that the Problem of Risk – contrary to the prevailing view in moral theory – is actually only a problem for very few, if any, existing moral theories.

Part 1.2 presents the Problem of Risk. **Part 1.3** discusses two representative responses that have been made on behalf of Moral Absolutism. **Part 1.4** sets out a choice-theoretic framework that formalises the Problem of Risk as an impossibility result for moral theories that have a particular structure. It also sets out a possibility result that shows how moral theories can potentially avoid the Problem of Risk. **Part 1.5** discusses whether any existing moral theories necessarily face the Problem of Risk. **Conclusion** follows.

1.2 Absolutism and the Problem of Risk

To illustrate the Problem of Risk, consider the following scenario:

Patient: You are a medical doctor in the emergency ward and you have a patient who needs urgent medical attention. Your options are: either give her Drug A or give her Drug B. The patient is either receptive to Drug A or receptive to Drug B. If the patient is A-Receptive, then giving Drug A will cure her, whereas giving Drug B will harm her. If, instead, the patient is B-Receptive, then Drug B will cure her and Drug A will harm her. You are uncertain whether the patient is A-Receptive or B-Receptive.[5]

Suppose that you are a Hippocratic absolutist: you believe that it is always morally impermissible to harm your patient. In cases of certainty, this moral norm gives you clear guidance about what to do: namely, don't harm your patient. In the present case, however, it is unclear which action is of the prohibited, harmful kind. As a Hippocratic absolutist, what are you to do, given your uncertainty?

Rights (Harvard University Press, 1990); McKerlie, "Rights and Risk." More recently, Sven Ove Hansson argues that a wide range of moral theories, from Consequentialism to Deontology, are unable to cope with risk and uncertainty. See: Hansson, "Ethical Criteria of Risk Acceptance"; Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World*. Michael Huemer argues that lexical priority theories – those that take some moral considerations to be more important than any number of 'lesser' considerations – also face the Problem of Risk. See: Huemer, "Lexical Priority and the Problem of Risk." Yoaav Isaacs (2014) presents the Problem for deontology, attempts solve it with a knowledge-first epistemology, then appears to disavow the solution. See: Isaacs, "Duty and Knowledge." Frank Jackson and Michael Smith (2006) raise the Problem for Moral Absolutism. See: Jackson and Smith, "Absolutist Moral Theories." They subsequently argue that the Problem extends to Moderate Deontology, which allows that, in extreme cases, deontological prohibitions (such as murder, or lying) can justifiably be violated so long as doing so would promote a sufficiently large amount of good. See: Lord and Maguire, "The Implementation Problem for Deontology."

5. This is a simplified version of a case given in: Jackson, "Decision-Theoretic Consequentialism."

To answer this, you need a *moral decision rule*: a rule that determines what you morally ought to do, given your options and your uncertainty. In their critique of absolutist moral theories, Jackson and Smith (2006) consider three candidate absolutist moral decision rules. However, they quickly show that each rule is highly problematic.

The first candidate rule, **Positive Probability**, prohibits any actions that have a *positive probability* of violating an absolutist moral norm. A problem with this rule is that, under our everyday conditions of uncertainty, it would prohibit all of our actions.[6]

The second candidate rule, **Certainty**, takes the opposite approach and encounters the opposite problem: it prohibits all and only those actions that are *certain* to violate a moral norm, but thereby permits practically all of our actions. The problem with this, of course, is that there seem to be many cases where an action is impermissible, even though it is not certain to violate a moral prohibition.

The third and final candidate rule that Jackson and Smith consider, **Threshold**, says: only perform actions that are less than t-probable to violate an absolutist moral norm. In cases like **Patient**, they consider a rule that says you ought to treat the patient if and only if doing so is less than t-probable to harm her. However, Jackson and Smith argue that this moral decision rule yields inconsistent verdicts. For example, consider the following case:

Second Patient: A second patient has just arrived, who presents the same symptoms and is just as likely as the first patient to be A-Receptive or B-Receptive. You must now decide whether to give the same drug to each patient, or to give them different drugs.

Suppose that for each patient the probability that your administering any particular drug will cause harm is less than t. As such, your moral decision rule holds that you ought to treat each patient, either with Drug A or Drug B.

However, suppose that if you treat both patients, the cumulative probability that at least one patient will be harmed is greater than t. According to your moral decision rule, you are therefore prohibited from administering the drug in both instances, even though each instance is required.

Thus, it seems that Threshold yields inconsistent verdicts about what you morally ought to do: you morally ought to treat each individual patient, but you ought not treat both. Jackson and Smith consider this an "especially implausible kind of moral

6. This is known as the 'Paralysis Problem'. See: Altham, "Ethics of Risk"; Hayenhjelm and Wolff, "The Moral Problem of Risk Impositions: A Survey of the Literature"; Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World*.

dilemma" and, on that basis, reject Threshold.[7] For want of any obvious alternative moral decision rule, but also having not exhaustively investigated all of the possible moral decision rules that absolutists might accept, they make the following conjecture: Moral Absolutism cannot give adequate moral guidance in cases involving risk.

If this conjecture is correct, then there is a strong case for rejecting Moral Absolutism. If we reject Moral Absolutism, then it turns out that no kinds of actions are always morally impermissible.

1.3 Responses to the Problem of Risk

Since Jackson and Smith did not prove the Problem of Risk, it has been open to defenders of Absolutism to offer proposals for avoiding it. I will review two such proposals below. As I shall argue, each of these proposals is problematic in its own right. However, to forestall ongoing debate, Section 1.4.4 proves that Jackson and Smith (2006) are correct: there is a particular type of absolutist moral theory that incorrectly determines the permissibility of risky actions. To the extent that defenders of Moral Absolutism are defending this type of theory, their responses fail.

1.3.1 A Rights-Based Approach

Ron Aboodi, Adi Borer, and David Enoch (2008) argue that a particular class of moral absolutist theories can adopt Threshold without generating inconsistent permissibility verdicts.[8] This class of theories is *patient-centred, rights-based, individualistic deontology*. A moral theory is 'patient-centred' when it grounds the permissibility of actions on their effects on other people; it is 'rights-based' when it holds that our moral duties are determined by the rights of others; it is 'individualistic' when it holds that only individuals – not groups – possess such rights. Many moral theorists appear to subscribe to this kind of view. As such, if it does avoid the Problem of Risk, then Jackson and Smith's critique is less important than previously thought.

In **Second Patient**, Aboodi et al.'s account holds that you ought to treat each patient, since each of your actions is not sufficiently probable to violate each patient's right against being harmed. Since it is permissible to treat each patient, it is

7. Jackson and Smith, "Absolutist Moral Theories," p. 276. For responses, see Ron Aboodi, Adi Borer, and David Enoch, "Deontology, Individualism, and Uncertainty: A Reply to Jackson and Smith," *Journal of Philosophy* 105, no. 2 (2008): 259–272; Patrick Hawley, "Moral Absolutism Defended," *Journal of Philosophy* 105, no. 5 (2008): 273–275; Lazar, "In Dubious Battle: Uncertainty and the Ethics of Killing"; Tenenbaum, "Action, Deontology, and Risk: Against the Multiplicative Model." At least some of these responses fail due to the impossibility result presented in Part 3.

8. Aboodi, Borer, and Enoch, "Deontology, Individualism"

permissible to treat both, even though the cumulative risk is above t. This is because there is no group agent consisting of 'Patient A-and-Patient B' that possesses the right against incurring a sufficiently high risk of harm. Thus, it seems that Jackson and Smith's critique is mistaken: there are absolutist moral theories that can give adequate advice in cases of uncertainty.

However, there are various problems with this particular approach.[9] More generally, there are problems with the general strategy of refining a theory's substantive moral commitments in order to avoid what is, as I shall argue, a structural problem. Simply put, the Problem of Risk does not concern *what* a moral theory believes to be morally important or *why*, but rather *how* it models the relations between these substantive elements.

In terms of their specific proposal, it is clear that – to the extent that they are defending Moral Absolutism – Aboodi et al. must deny the existence of group rights in *all* cases, since Threshold will give inconsistent verdicts when there are both group rights and individual rights at play. To avoid this inconsistency, they would need to abandon the absolutist idea that rights violations are always prohibited, or else explain why group rights can never come into conflict with individual rights.

Leaving this aside, there are still other commitments they must make, because it turns out that denying the existence of group rights is not sufficient to avoid inconsistencies. To see why, note that there are two ways of understanding patient-centred duties: on one reading, such duties are owed to particular, specified people (*de re*); on another reading, they are duties to people in general, whoever they happen to be (*de dicto*). On a *de re* reading, Aboodi et al.'s solution works.[10] If, however, Aboodi et al. opt for a *de dicto* reading, then their account gives the opposite verdict to the one they presented: your treating both patients would now be impermissible, because it has a sufficiently high probability of harming *someone* (whichever patient it happens to be).

This means that to retain their initial verdict, Aboodi et al. must not only deny group rights, but also adopt a strictly *de re*, patient-centred theory. However, this leads to other cases of inconsistency, calling for further substantive refinements. Consider the following case, based on Jackson and Smith (2016):[11]

Multiple Treatments: A third patient has arrived. She has three life-threatening

9. For other critical discussions of Aboodi et al.'s proposal, see: Lazar, "In Dubious Battle: Uncertainty and the Ethics of Killing"; Tenenbaum, "Action, Deontology, and Risk: Against the Multiplicative Model"

10. However, this opens them to other theoretical problems, such as the Non-Identity Problem: cases where it seems that we are entitled to choose avoidably harmful options, since the particular people we harm would not otherwise exist. See: Caspar Hare, "Voices from Another World: Must We Respect the Interests of People Who Do Not, and Will Never, Exist?," *Ethics* 117, no. 3 (2007)

11. Lord and Maguire, "The Implementation Problem for Deontology." This type of case is also credited to David Lewis, in McKerlie, "Rights and Risk," n. 3

conditions, each of which can be treated as follows:

Condition	Treatment
1	Drug A (if A-Receptive) or Drug B (if B-Receptive)
2	Drug C (if C-Receptive) or Drug D (if D-Receptive)
3	Drug E (if E-Receptive) or Drug F (if F-Receptive)

Suppose that there are no interaction effects between the drugs, but you are uncertain about the receptivity of the patient to particular drugs. If the probability that any particular drug will harm your patient is less than t, then Threshold permits you to administer any particular combination of drugs. At the same time, however, suppose that over the course of the three treatments, the cumulative probability of harm is greater than t. Assessing your actions in terms of individual treatments renders your actions permissible, whereas assessing them collectively renders them impermissible.

This case shows that denying the existence of group rights and committing ourselves to understanding duties *de re* as opposed to *de dicto* does not help us to avoid the inconsistency. What to do? Following Aboodi et al.'s approach, we could make further substantive commitments to ensure that we always get the right verdicts. For instance, given a particular metaphysical theory of action-individuation, we may determine whether your treating the patient is a single action (as in the coarse-grained description: 'Giving the A/D/E regime of drugs') or multiple actions (as in the fine-grained description: 'Give Drug A; Give Drug D; Give Drug E'). Adopting a coarse-grained or fine-grained approach to action-individuation may avoid this version of the dilemma. However, to some, making this kind of metaphysical commitment without any independent explanation would be objectionably *ad hoc*. Also, as it stands, it seems that we can expect still other cases of inconsistencies to arise, where aggregated and disaggregated moral considerations come into conflict, requiring further 'just-so' commitments.[12]

At best, the strategy of adopting further substantive refinements seems to be

12. For example, a proponent of this strategy must also take a stance on metaphysical questions about, for instance, personal identity: consider the metaphysical thesis called the Memory Criterion: a person is the same across time if and only if she can remember her previous experiences. Now, given a choice between imposing risk on a person within a short (memorable) time-frame or a long (not-memorable) time-frame, unless the deontologist decides on the validity of the Memory Criterion, it will turn out that imposing a series of risks actions across a long-time will be both permissible (because, accepting the criterion, the risks do not accumulate for any single person) and impermissible (because, denying the criterion, the risks do potentially accumulate for a single person). Likewise, Fission Cases – where psychological continuity of a person is 'branched' to multiple bodies, and some small quantum of risk is imposed on each – will generate further inconsistencies, unless the deontologist comes down on an answer, once and for all, about these metaphysical problems of personal identity. Needless to say, this seems like a laborious approach to avoiding the Problem of Risk.

an overcommittal, though potentially successful, approach to avoiding the Problem of Risk. At worst, this strategy risks having to play a never-ending game of philosophical 'Whack-A-Mole', with a structurally-identical problem re-emerging under increasingly complex permutations of substantive commitments. While it is important for a moral theory to be clear and consistent about its substantive commitments, it is worth considering whether the Problem of Risk can be solved by clarifying its structural commitments. To see this, we will next consider a different defence of Absolutism, one which also adopts the crucial, problematic structural assumption that generates the Problem of Risk.

1.3.2 A Sequential Decision Rule

Patrick Hawley (2008) argues that absolutists can avoid generating inconsistent verdicts by adopting the following sequential moral decision rule:

Eliminate: Do not perform actions that are greater than t probable to violate a prohibition.

Decide: Choose from the remaining options (perhaps by using some other decision rule).

Applying this approach to **Second Patient**, you should eliminate any option that has a higher than t probability of harming an individual (Hawley seems to assume that the moral duty is interpreted *de dicto*, not *de re*).[13] Since treating both patients is greater than t probable to lead to you harming a patient, you do not have the option of treating both. This means that you must treat only one of the patients. Hawley suggests performing an action that maximises expected utility.

Unfortunately, this proposal does not prevent implausible moral dilemmas arising. After all, in **Second Patient**, you do not have the option of treating only one patient. You must treat both, and by stipulation, whatever treatment you choose will be sufficiently low-risk for the particular recipient, but whatever combination you choose will be too risky and therefore impermissible.

It is important to note that this problem generalises: for any probability threshold t, if the number of states is greater than $1/(1-t)$, it will be possible that the threshold-based decision rule will consider each action permissible, but the aggregate of actions impermissible (and *vice versa*).[14] For instance, suppose that we continue using the given threshold of $t = 0.95$. This means that any action that is greater than 0.95 probable of violating an absolutist moral norm is impermissible. Suppose

13. Hawley, "Moral Absolutism Defended," p. 274

14. This is because Threshold, by its structure, commits Absolutism to moralised versions of the Lottery Paradox, as noted in Jackson and Smith, "Absolutist Moral Theories," n. 10.

that the relevant absolutist norms are: do not harm; do not use contraindicated drugs (these are drugs that are more likely than other drugs to harm the patient – using them would impose unnecessary risk on your patient).

Many Patients: Now there are thirty patients presenting the same symptoms. You must decide whether to treat them with the same drug (in which case, it would be Drug A or Drug B), or Drug A for some and Drug B for others. According to the blood tests, every patient is much more likely to be A-Receptive (0.9) than B-Receptive (0.1).

Given the information above, the probability that *all* of the patients are A-Receptive is only 0.04.[15] This means that treating all of the patients with Drug A is more than 0.95 probable to harm at least one of the patients. According to the Eliminate step of the sequential decision rule, you are prohibited from treating all patients with Drug A – indeed, it is "unavailable for rational deliberation".[16] The same applies, however, to the option of treating all patients with Drug B, since it is even less probable that all thirty patients are B-Receptive. However, you are also morally prohibited from treating some patients Drug A and some patients Drug B, since you are sure that Drug B is strongly contraindicated for each patient. This means that you face a moral dilemma. But this seems implausible: after all, if you have the best interests of each patient at heart, then you would give each patient Drug A.[17] Thus, it seems that Hawley's sequential decision rule does not address the underlying source of the Problem of Risk.

As we shall soon see, these responses to the Problem of Risk for Moral Absolutism fail because, in fact, there exists no adequate moral decision rule for the kinds of absolutist theories that Aboodi et al. and Hawley are purporting to defend. To show this, we must first formalise the Problem of Risk.

1.4 Formalising the Problem of Risk

So far, discussions of the Problem of Risk for Moral Absolutism have only been described it in informal terms. We have only been given a partial explanation of what is 'Moral Absolutism', what it means to 'give adequate moral guidance', and, indeed, what exactly are 'cases involving risk'. In this section, I will draw on the work of Dietrich and List (2017) to formalise the Problem of Risk.[18] This will allow

15. Calculated as: $0.9^{30} \approx 0.04$

16. Hawley, "Moral Absolutism Defended," n. 2

17. See also Caspar Hare, "Should We Wish Well to All?," *The Philosophical Review* 125, no. 4 (2016): 451–472

18. Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories."

us to determine, in more precise terms, which moral theories face the Problem of Risk and which ones avoid it.

1.4.1 A Choice-Theoretic Representation

The Problem of Risk applies to cases of moral decision-making in the face of uncertainty. To better understand such cases, we will define the following general formal framework.[19]

We will be concerned with what a moral theory says about the permissibility of an action in a given choice context. A *choice context* k consists of a set of actions, denoted $[k]$.[20] An action x is an element of $[k]$. In situations of moral decision-making under risk, actions are formally defined as functions from a set of *states* S to a subset of possible *outcomes* O.[21] In cases of risk, an action will have various potential outcomes, depending on which state of the world happens to obtain.[22] For instance, giving the patient Drug A in the state of the world in which she is A-Receptive will cure her, while doing so in the state of the world in which she is B-Receptive will not.[23] Note that outcomes need not be interpreted as *causal consequences*; they can equally be interpreted as fixing the *nature* of the action performed.[24] We will assume that we can measure the probability that the world is one way or another.[25] This probability could be interpreted subjectively (as representing our degrees of belief) or objectively (as evidential probabilities or chances). To simplify the analysis, we will assume that the possible states of the world are probabilistically independent of the action chosen: for instance, the fact that you give Drug A does not change the probability that the patient is A-Receptive.[26] Given a probability assignment to the possible states of the world, each of your actions has a *prospect*, which identifies the set of outcomes an action could bring about, as well as their respective probabilities.[27]

In moral decisions, we are concerned with the morally relevant *properties* P of an

19. For clarity and consistency, in my explanations of much of the formal framework below, I closely follow: ibid.

20. Let $\mathscr{K}$ be the set of all possible choice contexts. A choice context, k, is an element of $\mathscr{K}$. Within a choice context, you have a set of available actions that you can perform, denoted by $[k]$. $[k]$ is a subset of the universal set X of possible actions.

21. Savage, *The Foundations of Statistics*.

22. In a given choice context, there is a finite set of possible states of the world, $S = \{1, ..., n\}$, which are mutually exclusive and jointly exhaustive.

23. Formally, outcomes are action-state pairs, $O_{xn} = \langle x, S_n \rangle$, which represent an action having been performed in a particular state of the world.

24. Colyvan, Cox, and Steele, "Modelling the Moral Dimension of Decisions," p. 511.

25. We will assume that in a well-defined choice context, the probability of the states sums to 1.

26. For discussion, see: Jeffrey, *The Logic of Decision*, pp. 8-9.

27. Formally, an action's prospect is the set of its possible outcome-probability pairs in a context $\lambda(x, K) = \{\langle O_{x1}, \pi_1 \rangle, ..., \langle O_{xn}, \pi_n \rangle\}$. The universal set of choice contexts $\mathscr{K}$ includes the set of all possible combinations of actions and their prospects.

action in a particular context.[28] There are different kinds of properties that a moral theory might take to be relevant to the evaluation of an action's permissibility:

- A property P is an *option property* if its possession by the action in a particular context depends only on the action, not on the context.[29]
- A property P is a *context property* if its possession by an action in a particular context depends only on the context, not on the action.[30]
- A property P is a *relational property* if its possession by an action in a particular context depends on both the action and the context.[31]

Aside from identifying the set of morally relevant properties in a given choice context, a moral theory also tells us how an action's properties determine its permissibility. This will involve some kind of ranking of properties and sets of properties. As Dietrich and List note, different moral theories will adopt different types of rankings, depending on how they understand the relations between different types of moral properties.[32] For example, a ranking that is based on a *weighing relation* will presume that all moral properties are comparable – being either more, less, or equally weighty as, each other – leading to a transitive and complete ranking. This approach suits most consequentialist theories, but it is not fitting for some deontological theories.[33] A less committal ranking is one that is based on a *defeat relation*, which represents which sets of moral properties are more choice-worthy than others, while also allowing that the overall ordering of sets of moral properties may be incomplete or intransitive. In what follows, to ensure that we are not loading the deck against any particular moral theory, we will use a defeat relation rather than a weighing relation.[34]

28. Formally, a property P is a primitive object that picks out a set of action-context pairs $\langle x, k\rangle$, called the *extension* of P and denoted $[P]$. An action x in a particular context k *possesses* P when $\langle x, k\rangle \in [P]$. Note that the use of properties is a particularly important aspect of Dietrich and List's framework, allowing them to define a more general choice-theoretic framework than that of expected utility theory or other decision theories. See: Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories," p. 431.

29. Formally, for all x in X and all k, k' in $\mathscr{K}$, $\langle x, k\rangle \in [P]$ if and only if $\langle x, k'\rangle \in [P]$.

30. Formally, for all K in $\mathscr{K}$ and all x, x' in X, $\langle x, k\rangle \in [P]$ if and only if $\langle x', k\rangle \in [P]$.

31. Formally, such properties are those that are neither option properties nor context properties.

32. Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories."

33. For instance, deontological theories often allow that we have moral options to serve our own interests, even if doing so does not maximise the Good. In doing so, they deny that our interests are equally weighty as or more weighty than the Good we forego. See: Theron Pummer, "Whether and Where to Give," *Philosophy and Public Affairs* 44, no. 1 (2016): 77–95; Lazar, "Deontological Decision Theory and Agent-Centered Options."

34. A weighing relation or defeat relation, denoted $\trianglerighteq$, over sets of properties S is a binary relation whose relata are subsets of the universal set of properties, $\mathscr{P}$. When one set of properties S stands in this relation to another set S', formally $S \trianglerighteq S'$, then S weakly outweighs S', or S is ranked weakly above S', or S weakly defeats S'. Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories," p. 432.

The ranking of morally relevant properties entails a *moral decision rule*, which classifies actions as permissible or impermissible based on their respective bundles of morally relevant properties.[35] Using a ranking that is based on a defeat relation, a moral decision rule will hold that an action is permissible if and only if its set of properties is undefeated; an action is impermissible if and only if its set of properties is defeated by the set of properties of some other action.[36]

It is important to note that a moral decision *rule* is not a moral decision *procedure*. A *decision procedure* is a method for determining what ought to be done, consisting of a cognitive procedure, checklist or set of 'rules of thumb'. Decision procedures are adequate if they are reliable, accurate, efficient ways of determining what ought to be done. However, it is a *decision rule* that defines what ought to be done. Decision procedures thus aim to 'track' the verdicts of a decision rule. While both elements are important to moral decision-making, the Problem of Risk must be understood as the problem of determining what kinds of moral decision rules are logically compatible with a given background moral theory.[37]

1.4.2 Modelling Moral Absolutism

Moral Absolutism holds that there are some properties of actions that make that an action always morally impermissible. More precisely:

Moral Absolutism: Some kinds of actions have a particular wrong-making property $\bar{P}$, such that they are always impermissible.[38]

Although this definition helps to clarify the structure of Moral Absolutism, it needs to be further refined. This is because, as Jackson and Smith (2006) note, there is a trivial sense in which almost all moral theories are absolutist:

> Here we need to understand $\bar{P}$ as a property of an action as opposed to a relation between an action and available alternatives to that action. Classical

35. In Dietrich and List's terminology, a Moral Decision Rule corresponds to a Rightness Function.

36. Formally, where R is a function from k to $R(k)$ and $R(k)$ is a subset of $[k]$, a defeat relation is defined as: $R(k) = \{x \in [k] :$ it is not the case that the morally relevant properties of any action y in context k strictly defeat the properties of x in k. As Dietrich and List note (p. 468, n. 70), this could be further weakened to allow a moral decision rule to select the action that defeats the most alternatives, even though it is defeated by some.

37. On this distinction, see: R Eugene Bales, "Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?," *American Philosophical Quarterly* 8, no. 3 (1971): 257–265; Holly M. Smith, "Making Moral Decisions," *Noûs* 22, no. 1 (1988): 89–108.

38. Formally, if $\langle x, K\rangle \in [\bar{P}]$, then $\langle x, K\rangle \notin R(K)$. In terms of the defeat relation, actions that have the prohibited property are ranked lower than all actions that do not have that property. Also, to capture idea that such actions are always impermissible, the defeat relation must be nonreflexive, such that it does not rank $\bar{P}$ at least as highly as itself in cases where all of your available actions contain $\bar{P}$. Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories," p. 458.

> utilitarianism absolutely prohibits doing actions that fail to maximize utility. But an action's failing to maximize utility is a relation the action has to available alternatives. The distinctive feature of the kind of absolutism that we find, for example, in Kant and the Catholic tradition is that the absolutely prohibited kind is independent of the nature of any available alternatives.[39]

Following the earlier distinction between types of properties, we can capture the distinction between these kinds of Absolutism within the choice-theoretic framework as follows:

Option Absolutism: Whether an action in a particular context possesses the absolutely prohibited property $\bar{P}$ depends only on the action, not on the context.

Contrast Option Absolutism with:

Relational Absolutism: Whether an action in a particular context possesses the absolutely prohibited property $\bar{P}$ depends on both the action and the context.

Relational Absolutism can include comparative considerations in the definition of the $\bar{P}$. For instance, whether or not an action possesses $\bar{P}$ may depend on it having the least number of possible violations of a moral norm; or the lowest probability-weighted average value of norm violations; and so on.

1.4.3 Adequacy Conditions

As we saw earlier, some moral decision rules seem to be inadequate guides for risky situations. However, discussions of the adequacy conditions for a moral decision rule have tended to be elliptical and informal, leaving it unclear whether there could be some other moral decision rule that fits the bill. I will now identify and formalise a minimal set of adequacy conditions for moral decision rules that together generate an impossibility result for Option Absolutism.

To begin, the fact that a moral theory possesses a moral decision rule at all entails that it satisfies the following adequacy condition:

Action Guidance: For all choice contexts k and all actions x, x is either permissible or impermissible.[40]

This condition falls out of the fact that a moral decision rule is a function that maps an available action to either the subset of permissible actions or to the subset of impermissible ones. This may seem like an overly strong condition, since it entails

39. For notational consistency, here I have replaced Jackson and Smith's 'K' for $\bar{P}$. Jackson and Smith, "Absolutist Moral Theories," pp. 267-268. See also: Alan Gewirth, "Are There Any Absolute Rights?," *The Philosophical Quarterly* 31, no. 122 (1981): p. 4.

40. Formally, for all k in $\mathscr{K}$ and for all x in $[k]$, x is either in $R(k)$ or not in $R(k)$.

that a moral theory is never 'silent' about the permissibility of some action. After all, there are many cases where it seems that moral judgements simply do not apply (we do typically not expect moral theories to cast judgement on every trivial decision we make). We can alleviate this worry by noting that moral theories can satisfy Action Guidance by relying on a presumption of permissibility, such as: if a choice context has no morally relevant properties, then all of its actions are morally permissible.

A moral theory must do more than just satisfy Action Guidance. After all, moral decision rules like Certainty and Positive Probability satisfy Action Guidance: the former consistently regards almost all actions to be permissible, while the latter is consistent in prohibiting practically everything. These decision rules are inadequate because they conflict with other conditions. For example:

Impermissible Risk: There are contexts in which actions are impermissible even though they are not certain to violate a moral norm.

Permissible Risk: There are contexts in which actions are permissible even though they are not certain to conform to a moral norm.

There are various other adequacy conditions that we might put forward. To generate the impossibility result in Part 1.4.4, we will add just one other condition:

Dominance: An action x dominates another action y, if in a state-by-state comparison y does not defeat x in any state, but in some state, x defeats y. Given a choice between either a dominating action or a dominated one, a moral decision rule holds that you should perform the dominating action.

Dominance is a well-accepted condition for rational choice.[41] In cases where our actions affect others, it is also a compelling condition for moral choice. To illustrate using a slightly altered version of **Patient**, if under every contingency about the patient's receptivity, Drug A will lead to at least as good and potentially better outcomes for the patient than Drug B, then you ought to give Drug A and you ought not to give Drug B. Indeed, giving Drug B would needlessly deprive the patient of a potential cure – and in fact would do so in exchange for a heightened possibility of harm. In this way, Dominance is akin to a necessity constraint for risky situations, exhorting us to avoid choosing an action that imposes unnecessary risk by being guaranteed to be no better than – and potentially worse than – some alternative.

41. Here, as before, we are assuming Act-State Independence.

1.4.4 An Impossibility Result

In their discussion of Option Absolutism, Jackson and Smith were correct in holding that if such a moral theory rejects Certainty and Positive Probability, then it must adopt some kind of threshold approach. They illustrated this with the rule, Threshold. However, there is a wider class of threshold rules that could be adopted, depending on other statistical features of an action's prospect. For instance, a normative relevance function could instead take the *number* of possible moral norm violations to be relevant, or the *probability-weighted number*, or the *risk-weighted number*, or *mean-variance* value of possible outcomes, and so on.[42] These statistical features will not necessarily be functions of some fixed probability value; the same overall probability of violating a moral norm may not capture, for example, the number of norm violations or the range of such violations. Generalising from Threshold, any Option Absolutist theory that accepts Permissible Risk and Impermissible Risk must posit:

Prohibited Prospects The set of actions that have $\bar{P}$ are those that have a particular statistical feature P_s, where P_s applies to at least some actions that are not certain to violate an absolutist moral norm.

Option Absolutism must posit some kind of statistical feature that renders an action permissible or impermissible. A moral decision rule that satisfies Action Guidance maps actions to either the subset of permissible or impermissible actions (there being no 'in-between' deontic status). As such, this statistical feature will have a sharp cutoff, amounting to a statistical threshold of some kind. Option Absolutism must posit Prohibited Prospects or else it will be silent about the permissibility of risky actions, and will thereby fail all of the adequacy conditions set out earlier. This preliminary observation helps us to establish the following:

Proposition 1. *No Option Absolutist theory satisfies Impermissible Risk, Action Guidance, and Dominance.*

Proof. Suppose that an Option Absolutist theory satisfies Impermissible Risk. This implies that there are contexts k in which an action x possesses a prohibited prospect P_s (and, therefore $\bar{P}$). Now suppose that x is also in k', such that all other available actions y in k' are dominated by x. Since $\bar{P}$ is an option property, if x possesses that property in context k, then x also possesses it in context k'. That is, Option Absolutism holds that the presence of any action y does not change the impermissibility of x. By Action Guidance, this moral theory determines a moral decision

42. For a discussion of various ways of measuring and evaluating a prospect, see: Lara Buchak, *Risk and Rationality* (Oxford: Oxford University Press, 2013), chap. 2.

rule that entails that either: x is permissible in context k' (adhering to Dominance, but violating Option Absolutism) or x is not permissible in context k' (violating Dominance, but adhering to Option Absolutism).

□

To illustrate, suppose that in **Patient**, your Hippocratic moral theory holds that some treatments are 'off the table', being too dangerous to ever be permissibly used regardless of the alternatives. For instance, suppose that you should never give Drug C to your patient because it is too dangerous, even though there is some chance that it will harmlessly cure your patient. We can imagine cases where you could face a tragic choice between Giving Drug C or Giving Drug D, where Drug D *guarantees* death.[43] In such cases, Option Absolutism must maintain that Drug C is prohibited, even though it offers your patient the only means of survival.

How can Option Absolutists avoid this result? They may reject Dominance and argue that you face a prohibition dilemma, such that whatever you do, you are doing something you ought not to do. However, this seems highly implausible. Firstly, this is not the rather plausible kind of moral dilemma whereby there is a conflict of different, incommensurable considerations.[44] Here, the very same consideration is at play: namely, the well-being of your patient. This is also not the kind of moral dilemma where your two available actions (Drug C and Drug D) are equally bad or 'wrongful'. Indeed, Giving Drug D amounts to needlessly denying your patient the possibility of survival, whereas Giving Drug C grants your patient some prospect of survival.[45] Option Absolutism blindly rules out the one action that could save your patient.

Overall, it is difficult to see what could be said in defence of Option Absolutism in such cases. However, we may not have to say anything at all: in Part 1.5 we will see that the standard targets of the Problem of Risk can plausibly avail themselves of the following:

43. One might argue that for all such choice contexts, Do Nothing is always available and permissible, even if Do Nothing is also guaranteed to cause (by omission) your patient's death. The worry with this approach is that in cases of uncertainty, where 'doing' actions have some risk of being wrongful, we will always be required to 'do nothing'. See: Seth Lazar, "Deontological Decision Theory and the Grounds of Subjective Permissibility" (2018).

44. See Lisa Tessman, *Moral Failure: On the Impossible Demands of Morality* (Oxford: Oxford University Press, 2015).

45. Note that in this case, the commonly-held intuition that small differences in risk do not make a difference to the relative permissibility of actions is less compelling than normal. Even if the difference between Drug C and Drug D is only a small risk, it is nevertheless the case that there is nothing to say in favour of Drug D, but something to say in favour of Drug C, such that it would be wrong to not give Drug C instead of Drug D. *Pace* Tenenbaum, "Action, Deontology, and Risk: Against the Multiplicative Model," pp. 9-11.

1.4.5 A Possibility Result

Unlike Option Absolutism, Relational Absolutism avoids the Problem of Risk.

Proposition 2. *There are moral theories that satisfy Relational Absolutism, Permissible Risk, Impermissible Risk, Action Guidance, and Dominance.*

Proof. Given Permissible Risk and Impermissible Risk, there are contexts in which an action x is impermissible in k despite not being guaranteed to violate a moral norm. Suppose that the relevant relationally prohibited property is: $\bar{P}$ = "Having the lowest expected deontic value".[46] Given that $\bar{P}$ is a relational property, if x is in a context k' where there is only one other available action, y, which has a lower expected deontic value, then x does not have the relational property $\bar{P}$. Given Action Guidance, the moral decision rule determines the permissibility of actions x and y in k'.

Suppose for *reductio*, however, that this moral theory violates Dominance. This leads to an immediate contradiction: it is impossible that y dominates x, but y has lesser expected deontic value than x.[47] □

The decision rule given above is just one of the many available decision rules that satisfy the adequacy conditions we have set out. The availability of particular decision rules will depend on the particularities of the reasons structure, such as whether its defeat relation can be represented by a cardinally measurable value function (allowing for the moral decision rule: maximise expected deontic value, or a risk-weighted value function).[48] If the weighing relation can only be represented by an ordinal value function, then moral theories can, for example, attempt to adopt an Expected Borda Count.[49] If there is no probablistic information available at all, a moral theory could adopt a rule such as Maximin: choose the action that has the least worst

46. An action's 'expected deontic value' is the probability-weighted average of the deontic values of each of its possible outcomes, where the deontic value of an outcome represents (roughly) the importance of the moral norms at play in that outcome. This property is well-defined providing that a moral theory satisfies a set of axioms of expected value theory, suitably interpreted for moral theories rather than rational choice. For instance, see: Colyvan, Cox, and Steele, "Modelling the Moral Dimension of Decisions."

47. Given our background assumption of Act-State Independence.

48. Cardinal value functions represent whether one bundle of properties is more, less, or equally weighty than some other bundle, and also gives a meaningful measure of the difference in the weights of bundles. Formally, cardinal value functions are unique up to positive linear transformation (where the value function is multiplied by a positive number, plus a constant). On risk-weighted approaches to decision under risk, see: Buchak, *Risk and Rationality*.

49. An ordinal value function simply determines whether two bundles of properties are equally weighty or whether one is more weighty than another. Formally, such functions are unique up to positive transformation (that is, they maintain the same ordering when their values are multiplied by a positive constant). On the use of the Borda Count in a different theoretical context, see: William MacAskill, "Normative uncertainty as a voting problem," *Mind* 125, no. 500 (2016): 967–1004.

possible outcome.[50] In any case, once a moral theory accepts Relational Absolutism, it can potentially avoid the Problem of Risk.

1.5 Which Moral Theories Face the Problem of Risk?

The above results show that Option Absolutist theories cannot avoid the Problem of Risk, but Relational Absolutists can. How do these results bear on presently-held moral theories? I argue, firstly, that the standard targets of the Problem of Risk can actually be read as Relational Absolutists. I then show how Relational Absolutism can accommodate the kinds of considerations that would motivate one to accept Option Absolutism, namely: a concern for the inherent wrongfulness of particular kinds of actions, and a rejection of moral aggregation.

1.5.1 Will the Real Option Absolutists Please Stand Up?

It has often been suggested that simply subscribing to an *uncompromising* moral theory makes you subject to the Problem of Risk. As we saw earlier in this paper, moral theorists such as Kant, Anscombe, and Nozick have been tarred with the Absolutist brush, and subjected to the Problem of Risk.

However, despite their absolutist reputations, on closer inspection it actually appears that none of them are committed to Option Absolutism. I shall briefly argue that the above prototypical victims of the Problem of Risk can faithfully be interpreted as Relational Absolutists, not Option Absolutists.

There are tell-tale signs that a moral theory subscribes to Option Absolutism. For instance, a moral theory is Option Absolutist only if it accepts the possibility of prohibition dilemmas: cases where all of your actions are prohibited.[51] Also, if a moral theory is Option Absolutist, then it will ignore comparative information about actions. By contraposition, if a moral theory rejects the possibility of prohibition dilemmas or pays attention to comparative information, it is not an Option Absolutist theory.

As it turns out, Kant seems to hold that prohibition dilemmas are impossible.[52] Moreover, he appears to accept the relevance of comparative information about our actions, holding our obligating reasons (as Onora O'Neill calls them) can conflict

50. John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971).

51. Although the reverse is not true: accepting the existence of prohibition dilemmas does not commit you to Option Absolutism.

52. For a defence of this reading of his view, see: Onora O'Neill, "Instituting Principles: Between Duty and Action," *Southern Journal of Philosophy* 36, no. S1 (1998): 79–96.

and that, in such cases, we are required to choose the stronger reasons over the weaker ones:

> [A] subject may have, in a rule which he prescribes to himself, two obligating reasons (*rationes obligandi*), one or other of which is not sufficient to put him under obligation (*rationes obligandi non obligantes*), so that one of them is not a duty... When two such reasons conflict with each other, practical philosophy says, not that the stronger obligation takes precedence (*fortior obligatio vincit*) but that the stronger obligating reason prevails (*fortior obligandi ratio vincit*).[53]

Kant's position regarding obligating reasons can be modelled as a Relational Absolutist theory, whereby the prohibited actions are those that are not supported by the 'stronger' (or 'undefeated') obligating reasons:

- for every context k, the set of relevant properties is the set of obligating reasons P_R.
- the defeat relation ranks P_R over $P_{R'}$ if and only if P_R is a stronger obligating reason than $P_{R'}$.
- the class of absolutely prohibited actions $\bar{P}$ are those that have weaker obligating reasons than that of some other available alternative in that context.

The key suggestion here is that even though Kant is apparently committed to the idea that some kinds of actions are always morally impermissible, this alone does not commit him to Option Absolutism. It all depends on whether the relevant kinds are individuated with reference to one's other available actions. As it is, his discussions of conflicting obligating reasons and denial of prohibition dilemmas suggests that he is better understood as a Relational Absolutist. While this is by no means a definitive analysis, it should be enough to establish that purveyors of the Problem of Risk must do more than issue a summary judgement of whether a particular moral theory can deal with risky situations; they must show why such moral theories are specifically committed to Option Absolutism.

Let us now turn to another putative moral absolutist, G.E.M. Anscombe. At first glance, it is puzzling to label her a moral absolutist, since a headline message of her ethics that "the concepts of ... moral obligation and moral duty ... ought to be jettisoned."[54] It is hard to be a Moral Absolutist if you reject the importance of deontic verdicts.

53. Immanuel Kant, *The Metaphysics of Morals*, ed. Mary J Gregor (trans.) (Cambridge University Press, 1996), 6:224; O'Neill, "Instituting Principles: Between Duty and Action," p. 88.

54. G. E. M. Anscombe, "Modern Moral Philosophy," in *Virtue ethics*, ed. Roger Crisp and Michael A Slote (Oxford: Oxford University Press, 1997), 26–44.

However, in her discussions of the ethics of war, she does make such verdicts. Indeed, when she does so, she seems to commit herself to some version of Moral Absolutism:

> We may not commit any sin, however small, for the sake of any consideration, however great, and if the choice lies between our total destruction and the commission of sin, then we must choose to be destroyed.[55]

Although such statements certainly fit the caricature of Moral Absolutism, it is important to note that Anscombe's position regarding the relative importance of avoiding sinful actions versus ensuring our survival does not commit her to Option Absolutism. To determine whether Anscombe is an Option Absolutist, we must determine whether she takes the relative merits of our other available actions to determine whether an action is impermissible (or 'sinful').

In fact, even Anscombe appears to adopt a Relational Absolutist approach to permissibility. For example, in her discussion of President Truman's decision to bomb Nagasaki, she argues that the decision was 'murderous' because it was marred by "the fixation on unconditional surrender [and] the disregard of the fact that the Japanese were desirous of negotiating peace."[56] Bombing Nagasaki was murder, according to Anscombe, because there were better options available, not because killing the innocent is a categorically prohibited action. Indeed, she makes this last point explicit:

> I intend my formulation to be taken strictly; each term in it is necessary. For killing the innocent, even when you know as a matter of statistical certainty that the things you do involve it, is not necessarily murder. I mean that if you attack a lot of military targets, such as munitions factories and naval dockyards, as carefully as you can, you will be certain to kill a number of innocent people; but that is not murder. On the other hand, *unscrupulousness in considering the possibilities turns it into murder.*[57]

Thus, in so far as Anscombe can be read as making absolutist 'moral ought' claims at all, she seems to accept that whether an action possesses a particular absolutely prohibited property (say, that of being 'murderous') at least partly depends on how that action compares to its alternatives. This suggests that Anscombe's ethical theory cannot be easily categorised as a version of Option Absolutism. Rather, a charitable reading of her view will be that of Relational Absolutism, allowing it to avoid the Problem of Risk.

Finally, let us consider whether Robert Nozick's Libertarianism, and rights-based theories in general, necessarily face the Problem of Risk. In fact, it seems that even

55. G. E. M. Anscombe, *Ethics, Religion and Politics* (Oxford: Basil Blackwell, 1981), p. 79.
56. ibid., p. 65.
57. ibid., p. 66 Emphasis mine.

Nozick believes that rights-based theories have problems dealing with risk:[58]

> Actions that risk crossing another's boundary pose serious problems for a natural-rights position ... Imposing how slight a probability of a harm that violates someone's rights also violates his rights? ... [Solutions relying on expected value calculations] cannot be utilized by a tradition which holds that stealing a penny or a pin or anything from someone violates his rights. That tradition does not select a threshold measure of harm as a lower limit, in the case of harms certain to occur. It is difficult to imagine a principled way in which the natural-rights tradition can draw the line to fix which probabilities impose unacceptably great risks upon others.[59]

It is arguable that Nozick is overstating the incompatibility of rights-based theories with expected utility calculations. Notice that in the often-quoted passage above, Nozick is concerned with finding a principled means of determining which kinds of risks count as rights-violating. Under Dietrich and List's framework, this amounts to determining which kinds of risks are included in the set of morally relevant properties. Nozick seems to be worried that if we have rights against risk being imposed on us, then all of our actions will be impermissible because all of our actions tend to impose some risk on others.[60] However, this worry is misplaced.

Simply put, a rights-based theory can accept that all risk impositions matter, while holding that they do not all matter equally.[61] The key question will be whether and, if so, under what conditions, our rights against risk may be outweighed by other moral considerations. Short-cutting some lengthy Nozickian exegesis, his theory corresponds to roughly the following reasons-structure:

- for every context k, the set of morally relevant properties consists of rights-based considerations P_R and goodness-based considerations P_G.
- the defeat relation ranks P_R over P_G for all instances of P_R and P_G.[62]
- within the class of P_R actions, those that involve *you* violating a right P_{R_i} are ranked lower than those that involve *others* violating a right $P_{R_{\neg i}}$.

This reasons structure is consistent with Relational Absolutism. For instance, Nozick's theory can adopt the following moral decision rule:

58. Nozick, *Anarchy, State, and Utopia*. See also: McKerlie, "Rights and Risk"; Altham, "Ethics of Risk."

59. Nozick, *Anarchy, State, and Utopia*, pp. 74-5

60. Many have echoed this worry. See: Hayenhjelm and Wolff, "The Moral Problem of Risk Impositions: A Survey of the Literature"; Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World*.

61. On this approach, see: McCarthy, "Rights, Explanation, and Risks."

62. This may not be quite right: in a well-known footnote, Nozick leaves open that there may be extreme cases where a right may be permissibly violated for the sake of the good. Nozick, *Anarchy, State, and Utopia*, pp. 29-30.

- the class of absolutely prohibited actions $\bar{P}$ are those in which you do not maximise your expected degree of rights conformity.[63]

Overall, it is both uncharitable and inaccurate to assume that uncompromising theories (like those of Nozick, Anscombe and Kant) are Option Absolutist theories. Rather, they are instances of a class of moral theories called *lexical priority theories*: those that hold that some kinds of moral considerations cannot be defeated by any number of particular other considerations. In general, lexical priority theories can accept Relational Absolutism by, for example, holding that we should choose the action that best upholds the higher consideration, using only the lower considerations to break ties between actions.[64] Hence, these theorists can avoid the Problem of Risk.

1.5.2 Objections to Relational Absolutism

Notwithstanding the fact that Option Absolutism can lead to violations of Dominance, some moral theorists may nevertheless baulk at accepting Relational Absolutism. And understandably so: it seems, at least at first glance, that Relational Absolutism will misrepresent moral theories that take some actions to always be wrongful or those that prohibit the weighing and aggregation of people's lives. However, I will suggest that such theories can accept Relational Absolutism without giving up these important substantive commitments.[65]

One motivating concern is based on the idea that there is a close connection between an action's permissibility and whether it is rightful or wrongful. From this perspective, accepting Relational Absolutism seems to ignore the idea that some actions are always morally *wrongful*, and therefore cannot be morally *permissible*. For instance, it seems bizarre to say that it is *morally permissible* to play five-bullet Russian Roulette on an unconsenting individual just because the only available alternative is to play six-bullet Russian Roulette. Both actions are, in a very immediate sense, morally *impermissible*. If that means giving up Dominance, so be it! Many moral theorists – deontologists, in particular – seem to share this kind of sentiment. However, I will briefly sketch how this position can be captured by Relational Absolutism, allowing one to maintain Dominance while still recognising the wrongfulness of particular kinds of actions.

63. One might object that any such rule will be 'goal-directed', and hence ruled out by Nozick: ibid., pp. 30-33. Such a charge would need to show why the moral decision rule given here necessarily commits us to treating others as mere means to our ends, which is Nozick's chief concern. However, so long as we are careful about the reasons structure, however, it doesn't seem like there will be any such commitment.

64. See: Chapter 2 and Chapter 3.

65. Thanks to Seth Lazar for pressing these points.

Consider the following view: some kinds of actions, such as playing Russian Roulette on someone – are always wrongful. Performing such an action provides grounds for shame, compensation, or perhaps even punishment. However, if you find yourself forced to play a round of Russian Roulette on someone with five bullets or six bullets, although both options are wrongful, you ought to choose to play with five bullets, since it is the *least* wrongful action available (it at least gives your victim *some* chance of survival).

I submit that this would be a Relational Absolutist view, whereby the relevant properties of an action include wrongfulness properties, and the absolutely prohibited actions are those that do not minimise wrongfulness. Of course, formally speaking, your moral decision rule will deem your five-bullet option 'permissible', but this is a very thin sense of permissibility (after all, your action is still wrongful). Moreover, this permissibility verdict is extremely *modally fragile* – under almost any other choice context, your action would be clearly impermissible. Note also that accepting Relational Absolutism does not rule out the possibility of strong moral dilemmas: it only rules out such dilemmas when one action dominates another. In other cases where Dominance does not apply, there may be a prohibition dilemma at hand. For instance, if a moral theory holds that the impermissible actions are those that are defeated by some other action, then they may encounter prohibition dilemmas in cases where all actions are defeated by some other action. Overall, accepting that some actions can be inherently wrongful and that prohibition dilemmas exist does not confine a moral theorist to Option Absolutism. Such theorists can accept an appropriately interpreted Relational Absolutist approach, and thereby avoid the Problem of Risk.

A second reason for resisting Relational Absolutism is based on a rejection of particular kinds of moral aggregation. On this kind of view, individual claims should not be weighed against each other – the numbers, as Taurek put it, don't count.[66] To weigh the aggregate interests of a collection of people against those of others would be to overlook the separateness of persons.[67] The worry is that accepting Relational Absolutism and Dominance commits us to violating this tenet.

However, this does not follow. Relational Absolutism can deem the numbers of claims to be morally irrelevant, and instead rank actions according to whether or not they give individuals the appropriate chance to have their claim satisfied. In terms of the defeat relation, the highest-ranked properties of actions are those that

66. John M. Taurek, "Should the Numbers Count?," *Philosophy and Public Affairs* 6, no. 4 (1977): 293–316.

67. Thomas Nagel, *Mortal Questions* (Cambridge University Press, 1979); John Rawls, *Justice as Fairness: A Restatement*, ed. Erin Kelly (Cambridge, Massachusetts and London, England: Harvard University Press, 2001); T.M. Scanlon, *What We Owe to Each Other* (Cambridge, Massachusetts and London, England: The Belknap Press of Harvard University Press, 1998); Alex Voorhoeve, "How Should We Aggregate Competing Claims?," *Ethics* 125, no. 1 (2014): 64–87.

give the most appropriate probabilities of satisfying individuals' respective claims. This is compatible with Dominance, since it only requires that you do not choose an action that has an unnecessary risk of satisfying no-one's claim. More generally, accepting Relational Absolutism does not entail accepting any particular stance towards moral aggregation. Thus, even ardent anti-aggregationists can accept this version of Relational Absolutism.

Overall, it seems that few moral theories, if any, are necessarily Option Absolutist. It also seems that there are few reasons, if any, to be an Option Absolutist rather than a Relational Absolutist. For critics of a moral theory to successfully 'use structure as a weapon', their best approach may instead be to show that their target moral theories face the Problem of Risk under stronger adequacy conditions for moral decision rules.[68] However, this remains to be seen.

1.6 Conclusion

The Problem of Risk has been levelled against a number of moral theories over the years. It has led to valuable improvements in our understanding of moral theories. This paper has demonstrated, however, that much of the discussion of the Problem of Risk has been mistaken. Jackson and Smith (2006) are correct that Option Absolutism faces the Problem of Risk. However, subsequent attempts to extend the Problem of Risk to other moral theories fail, since such theories can be interpreted as accepting Relational Absolutism. In fact, it is arguable that even prototypical Option Absolutists are actually Relational Absolutists. The upshot is that since few, if any, moral theories are committed to Option Absolutism, it turns out that a much wider range of moral theories can potentially guide us through uncertain situations than previously thought possible.

68. To use Nozick's phrase: Robert Nozick, "Moral Complications and Moral Structures," *Natural Law Forum* 13 (1968): p. 44.

Part II

Solutions

Chapter 2

Moral Priorities Under Risk

2.1 Introduction

A public official is deliberating about whether to approve a marginal increase in the speed limit for autonomous vehicles. Surveys show that this increase would improve passenger satisfaction, but would not lead to any other substantial improvements. Against this move, however, is evidence that increasing the speed limit poses some risk of increasing the current incidence of pedestrian deaths: some people crossing the road may underestimate the speed of the self-driven vehicles as they quietly shuttle through the streets. There are no other relevant considerations. The decision is hers alone.

The official takes a moment to reflect on the moral considerations at hand. One such consideration is public safety: she ought to choose the option that minimises pedestrian deaths. The other consideration is people's pleasure: increasing passenger satisfaction will increase the amount of pleasure in the world. Although both considerations are morally significant, to her mind they are not equally important. In fact, she believes that no amount of passenger satisfaction could ever morally justify a pedestrian's death. As far as she is concerned, considerations of public safety are the primary moral consideration at hand.

The primacy of public safety in this case can be spelled out more precisely using the notion of *lexical priority*.[1] A moral consideration has lexical priority over another just in case, given the choice between upholding a higher-ranked consideration versus upholding any number of lower-ranked ones, we ought to uphold the higher-ranked consideration. Lower-ranked considerations are only relevant insofar as they help to break ties between options that uphold higher-ranked considerations equally well.

1. On the history and applicability of this concept to various domains in moral and political philosophy, see: John Rawls, *A Theory of Justice: Revised Edition* (Cambridge, MA: Belknap Press, 1999). Rawls himself takes lexical priorities to only be a 'useful approximation' for the purposes of his theory.

Lexical priority theories thus prohibit trade-offs between different kinds of moral considerations. In the official's case, considerations of passenger satisfaction ought not to be traded-off against pedestrians' lives.

To her disappointment, however, the official realises that her awareness of these moral considerations is not sufficient to guide her decision. The problem, of course, is that she does not know whether increasing the speed limit will in fact lead to more deaths. Since she is to some degree uncertain about whether her choices will lead her to violate or uphold particular moral considerations, she faces a *moral decision under risk*. What she requires is a *moral decision rule*: a rule that identifies which options are morally permissible, given her degrees of uncertainty. The question is: is there a moral decision rule that both maintains the lexical priority of public safety over additional passenger satisfaction, while also giving acceptable guidance in moral decisions under risk?

If the recent decision-theoretic critique of lexical priorities is correct, then it turns out that there is no acceptable moral decision rule available to the official.[2] This is because lexical priority theories appear to commit a decision-maker to at least one of the following problems: the Permissiveness Problem, the Low Risk Problem, and the Agglomeration Problem. The Permissiveness Problem consists of cases where all of one's options are equally permissible, simply because they have some positive probability of violating a lexical priority. The Low Risk Problem arises when arbitrarily small probabilities make particular actions impermissible. The Agglomeration Problem involves cases where lexical priority theories give contradictory advice depending on whether the relevant moral considerations are viewed separately or together. Lexical priority theories must avoid each of these problems in order to have an acceptable moral decision rule for risky situations.

This chapter uses lexicographic expected value theory to rebut the decision-theoretic critique. As I shall argue, many of the problems that have been raised for lexical priority theories are actually the result of inappropriate decision-theoretic modelling. Once a more appropriate model is applied, the problems do not arise. Nevertheless, although the decision-theoretic critique fails in this respect, it succeeds in revealing important but under-theorised aspects of lexical priority theories.

Part 2.2 briefly sets out the preliminaries for modelling a moral theory using decision theory. **Part 2.3** responds to the Permissiveness Problem. **Part 2.4** responds the Low Risk Problem. **Part 2.5** responds to the Agglomeration Problem. **Part 2.6** discusses the value and limits of decision-theoretic critiques of moral theories, and identifies further questions that lexical priority theories must answer to give a complete account of moral decision-making under risk. **Conclusion** follows.

2. Huemer, "Lexical Priority and the Problem of Risk"; Jackson and Smith, "Absolutist Moral Theories"; Lord and Maguire, "The Implementation Problem for Deontology."

2.2 Preliminaries

The decision-theoretic critique of lexical priority theories takes place on the following theoretical terrain: firstly, it proceeds on the assumption that *expected value theory* is an appropriate framework for representing moral theories; secondly, the objections to lexical priority views are premised on a *particular way of modelling* lexical priorities using expected value theory. I will explain these theoretical points in turn.

In its normative application, expected value theory determines what you ought to do, given your uncertainty. It identifies the best option relative to a description of a decision problem.[3] A decision problem consists of: the *options* available to you; the possible *states* of the world; the *probabilities* of those states (represented by some value in the [0, 1] interval), given that you perform some option;[4] and the value of the *outcomes* that result from choosing a particular option in a given state of the world. According to expected value theory, the best option is the one that maximises *expected value*, where this is the option whose possible outcomes together have the greatest sum of probability-weighted value.

To apply expected value theory to moral theory, we must assume that the relative importance of moral considerations can be numerically represented by a value function.[5] This 'moral' value function must assign equal value to equally important considerations, greater value to more important considerations, and lesser value to less important considerations.[6] However, to operate in a context of risk, a moral value function must also accurately represent the *differences* in the relative importance of moral considerations. That is, we must assume that the importance of moral considerations can be represented by a cardinal moral value function.

As we shall see, the precise specification of the cardinal value function is important to how a moral theory operates in an expected value framework. A common modelling assumption in discussions of lexical priority theories is that higher-ranked

3. For an accessible introduction to the normative applications of expected utility theory, see: R.A. Briggs, "Normative Theories of Rational Choice: Expected Utility Theory," *The Stanford Encyclopedia of Philosophy*, 2017,

4. Here I wish to remain neutral about whether evidential decision theory or causal decision theory is correct. Readers who prefer evidential decision theory can read the probabilistic dependence of states as conditional probabilities (the probability of a particular state, given that a particular option is chosen), whereas those who prefer causal decision theory can read the dependence in terms of the probabilities of subjunctive conditionals, imaging functions, dependency hypotheses, etc. See: James M. Joyce, *The Foundations of Causal Decision Theory* (Cambridge; New York: Cambridge University Press, 1999), chap. 5

5. More precisely, in an expected value framework, the values are assigned to the *outcomes* in which the moral considerations are upheld or violated. For ease of discussion, however, I will simply say that the considerations have value.

6. Slightly more formally, for all considerations C_1 and C_2, a moral importance relation w (where $\succ_w$ stands for 'more important than' and $\sim_w$ stands for 'equally as important as'), and a moral value function v: $C_1 \succ_w C_2$ if and only if $v(C_1) > v(C_2)$; and $C_1 \sim_w C_2$ if and only if $v(C_1) \sim v(C_2)$.

considerations are *infinitely* more important than lower-ranked considerations, and so should be represented by an infinite value difference.[7] As we shall see, this is a problematic and unnecessary assumption. Strictly speaking, lexical priority theories only subscribe to an ordinal ranking of moral considerations: given a choice between upholding either a higher-ranked consideration or any number of lower-ranked considerations, it is *more* important to uphold the higher-ranked consideration. Lexical priority theories need not make any commitment about *how much more* important it is to uphold a higher-ranked consideration than lower-ranked ones. Infinite values simply offer one way of representing the fact that lexical priority theories stratify some kinds of moral considerations over others.

To test whether lexical priority theories can give adequate guidance in cases of risk, we will model the official's case as follows. We will assume that she has only two available options: increase the speed limit (by some fixed amount) or maintain the current speed limit. There will be only two possible and mutually exclusive states of the world: either the pedestrians will be careful when crossing the road (*Careful*) or they will not (*Careless*). We will allow that the probabilities of these states may be affected by the option chosen. Following the critics of lexical priority theories, we will hold that outcomes in which a lexical priority is violated – for instance, cases where there is an increase in pedestrian deaths – have infinite disvalue ($-\infty$), whereas outcomes in which a lexical priority is upheld are assigned some finite value (f). The above framework sets out a particular way of determining what lexical priority theories require of us in cases of uncertainty. One might object to this framework on a number of fronts.[8] For dialectical purposes, however, we will assume that it is appropriate. We will see that by making amendments within the expected value framework, lexical priority theories can satisfactorily deal with the decision-theoretic critique.

2.3 The Permissiveness Problem

Lexical priority theories hold that we are morally prohibited from trading-off higher considerations for the sake of lesser ones. Framing lexical priority theories in this way, an important question is: what are we to do when all of our options have some

7. See, for example: Jackson and Smith, "Absolutist Moral Theories"; Lord and Maguire, "The Implementation Problem for Deontology"; Colyvan, Cox, and Steele, "Modelling the Moral Dimension of Decisions"; Huemer, "Lexical Priority and the Problem of Risk"; Hayenhjelm and Wolff, "The Moral Problem of Risk Impositions: A Survey of the Literature"; Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World*; Adam Bjorndahl, Alex John London, and Kevin J. S. Zollman, "Kantian Decision Making Under Uncertainty: Dignity, Price, and Consistency," *Philosopher's Imprint* 17, no. 7 (2017): 1–22. Huemer's model is an informal rendition of the infinite values model introduced in Jackson and Smith, "Absolutist Moral Theories."

8. Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, chap. 8

positive probability of being of this prohibited kind? Some worry that lexical priority theories may commit us to a life of implausibly strong moral dilemmas, whereby all of our options are prohibited, merely by virtue of having a positive probability of violating a lexical priority. They call this the Paralysis Problem.[9]

However, when the problem is cast in an expected value framework, the Paralysis Problem is in fact better understood as the Permissiveness Problem. Expected value theory, after all, exhorts us to choose the option that maximises expected value, even if that option has low (or, indeed, negatively infinite!) expected value. As we shall see, when lexical priority theories are modelled using infinite values, the problem is that all of our options turn out to be *equally permissible* simply by virtue of having some positive probability of violating a lexical priority.

To illustrate, suppose that the official must decide on the basis of the following evidence: one report suggests that increasing the speed limit may greatly increase pedestrian deaths, whereas another report suggests that maintaining the current speed limit may also increase pedestrian deaths, albeit to a lesser extent. Given that both options are risky, what ought she to do? As shown in Table 2.1, since both options have some positive probability of violating a lexical priority, both have negatively infinite expected value. According to the expected value model, both options are therefore permissible because they have the same expected value.

Action	*Careless* p	*Careful* $(1-p)$	*Expected Value*
Increase speed limit	$-\infty$	f_1	$-\infty$
Maintain speed limit	$-\infty$	f_2	$-\infty$

Table 2.1: The Permissiveness Problem.

Clearly, however, this can lead to absurd results. For example, even if one option is far riskier than the other, it will have the same expected value and therefore will be deemed equally choice-worthy. The Permissiveness Problem seems to show that lexical priority theories, blinded by their own fanaticism, lack the ability to distinguish impermissible risky options from permissible ones.

As it turns out, however, the Permissiveness Problem does not pose a serious challenge to lexical priority theories. Rather, the problem is due to a flawed decision-theoretic model of lexical priority theories. Specifically, the Permissiveness Problem arises because infinite values swamp expected value calculations, rendering all options equally permissible regardless of their riskiness. A more appropriate model

9. For versions of this problem, see: Huemer, "Lexical Priority and the Problem of Risk," p. 337; Bjorndahl, London, and Zollman, "Kantian Decision Making Under Uncertainty: Dignity, Price, and Consistency," p. 8; Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World*, chap. 2; Hayenhjelm and Wolff, "The Moral Problem of Risk Impositions: A Survey of the Literature," E26-51.

will allow lexical priority theories to use expected value calculations to distinguish between permissible and impermissible options. This gives us good reason to look for non-infinitistic, expected value representations of lexical priority theories under risk.

The rich literature on Pascal's wager offers many candidate approaches.[10] One example is lexicographic expected value theory. Unlike most other approaches, the lexicographic approach constitutes a minimal departure from orthodox expected value theory.[11] Indeed, it is simply a generalisation of standard expected value theory. Where standard expected value theory operates with a one-dimensional value function, the lexicographic theory operates with a multi-dimensional function. It also has the benefit of being well known and comprehensively theorised in contemporary economic and decision theory.[12] Famously, Rawls defended a version of lexicographic decision theory in his Theory of Justice, although he did not apply it to decision-making under risk (due to his denial that probabilities are available behind the veil of ignorance).[13] As we shall see, the lexicographic model allows lexical priority theories to systematically avoid the Permissiveness Problem.

Using the lexicographic model, rather than using infinite values, we will represent lexical priorities using a ranking of finite-valued value functions, $< v_1, v_2, \ldots, v_n >$. In the official's case, we have assumed for simplicity that there are only two kinds of moral considerations at hand: public safety and additional passenger satisfaction. As such, we will require only two value functions, $< v_1, v_2 >$, where v_1 represents the moral significance of the level of public safety in that outcome and v_2 represents the moral significance of the level of passenger satisfaction in that outcome. We represent the priority of public safety over additional passenger satisfaction using the following moral decision rule: the official ought to choose the act that has the highest expected value for v_1 and – just in case there is a tie among the options – then she ought to choose the act that has the highest expected value for v_2. In the unlikely event that there is yet another tie, she is permitted to choose either option

10. For a review of such approaches, see: Alan Hájek, "Waging War on Pascal's Wager," *The Philosophical Review* 112, no. 1 (2003): 27–56

11. The Permissiveness Problem can also be avoided without abandoning infinite values, using Relative Utility Theory – see: Paul Bartha, "Taking Stock of Infinite Value: Pascal's Wager and Relative Utilities," *Synthese* 154, no. 1 (January 2007): 5–52. For simplicity of exposition, I have chosen an alternative, lexicographic approach.

12. See, for example: Nicolas Houy and Koichi Tadenuma, "Lexicographic compositions of multiple criteria for decision making," *Journal of Economic Theory* 144, no. 4 (2009): 1770–1782.

13. Rawls's theory adopts a lexicographic framework. However, his approach rejects expected utility maximisation altogether, in favour of a maximin decision rule for decision-making under ignorance (John Rawls, "Some Reasons for the Maximin Criterion," *The American Economic Review* 64, no. 2 (1974): 141–146. An early exposition of the idea of lexicographic utility functions is found in Von Neumann and Morgenstern, *Theory of Games and Economic Behavior*. However, it wasn't until later that lexicographic utility functions received their first systematic study in Melvin Hausner, "Multidimensional Utilities," chap. 12 in *Decision Processes*, ed. R. M. Thrall, C. H. Coombs, and R. L. Davis (New York: John Wiley / Sons, Inc., 1954).

(since there are no other considerations in this example).[14]

Without loss of generality, we will assume that the official has a confidence of 0.95 that the public will be careful and we will assign more-or-less arbitrary lexicographic values to the various outcomes. This will give us a clear demonstration that our chosen moral decision rule can guide decision-making in cases where all options are risky.

Action	*Careless*	*Careful*	*Expected Value*
Increase speed limit	$-1000, 1$ $p = 0.05$	$0, 1$ $(1 - p) = 0.95$	$-50, 1$
Maintain speed limit	$-2, -1$ $q = 0.05$	$0, -1$ $(1 - q) = 0.95$	$-0.1, -1$

Table 2.2: Lexicographic Expected Value.

As shown in Table 2.2, the option that uniquely maximises expected moral value in this case is to maintain the speed limit. However, that is not the important point. The important point is that there is an option that uniquely maximises expected moral value, even though all options have some probability of violating a lexical priority. As the official is not permitted to pursue just any of her available options, the Permissiveness Problem has been avoided. This shows that infinite values were indeed at the root of the problem. By eschewing infinite values, the expectations of the official's options are now sensitive to, among other things, the probabilities of the states. The lexicographic model is therefore a more appropriate decision-theoretic representation of lexical priority theories.

2.4 The Low Risk Problem

Although lexicographic decision theory solves the Permissiveness Problem, it seems to expose lexical priority theories to the Low Risk Problem. This involves cases where an intuitively acceptable option is ruled out simply because it has a positive – albeit negligible – probability of violating a lexical priority. To many, such cases give us grounds for rejecting lexical priority theories altogether.[15]

To illustrate the Low Risk Problem, let us suppose that the official has reason to believe that increasing the speed limit will almost certainly cause the public to become more careful. By contrast, maintaining the current speed limit will encourage

14. Slightly more precisely, for two options $A = (a_1, a_2)$ and $B = (b_1, b_2)$: $A \sim B$ if and only if $[(a_1 = b_1)\&(a_2 = b_2)]$; $A \succ B$ if and only if $(a_1 > b_1)$ or $[(a_1 = b_1)\&(a_2 > b_2)]$. This generalises to n-components.

15. For example, see: Huemer, "Lexical Priority and the Problem of Risk."

complacency, leading to an almost equal probability of carelessness or carefulness. According to the lexicographic model, what ought she to do?

Action	*Careless*	*Careful*	*Expected Value*
Increase speed limit	$-1000, 1$ $p = 0.001$	$0, 1$ $(1 - p) = 0.999$	$-1, 1$
Maintain speed limit	$-2, -1$ $q = 0.45$	$0, -1$ $(1 - q) = 0.55$	$-0.9, -1$

Table 2.3: The Low Risk Problem.

As shown in Table 2.3, the official maximises expected moral value by maintaining the speed limit, despite the fact that it is very improbable that the public would be careless, were she to increase the speed limit. Critics of lexical priority theories hold that in these types of cases, there is a point at which the probabilities of violating a lexical priority are so low as to be negligible. It cannot be that *any* prospect of upholding a lexically-prior consideration, no matter how slim, is more important than a certainty of upholding *any* number of lesser considerations.

To appease proponents of the Low Risk Problem, lexical priority theories must give a principled explanation for why decision-makers should ignore very unlikely prospects of upholding a lexically-prior consideration. One approach is to posit a probabilistic threshold, t, that governs which possibilities decision-makers should ignore and which they should attend to. Note that this option is available because expected value theory merely identifies the best option relative to a specification of the decision problem. It is silent with respect to what information is or is not included in the decision problem. It is therefore open to lexical priority theories to supplement the lexicographic decision model with rules for determining which possibilities are relevant to a decision and which should be ignored.

Critics of lexical priority theories doubt that there is any principled way of fixing a value for t.[16] However, they seem to assume that the threshold value must be context invariant: that is, the same under all circumstances.

Admittedly, it is difficult to see what could justify a universal choice of t. However, since there is no need to assume context invariance, I will offer a context-variant approach to assigning a value to t. On this approach, the chosen threshold should be that which, when followed as a rule, maximises expected moral value.[17]

16. Jackson and Smith, "Absolutist Moral Theories," p. 276; Isaacs, "Duty and Knowledge," p. 97.

17. A similar justification is given in non-moral cases by Hannes Leitgeb, "The Stability Theory of Belief," *Philosophical Review* 123, no. 2 (2014): pp. 150-151. Note that the justification put forward here for ignoring low probabilities differs from that proposed in Smith, "The Subjective Moral Duty to Inform Oneself before Acting," in two main respects. Firstly where Smith posits that practical norms are 'tolerant' of slight deviations from infinite precision, the approach put forward

For example, a policy of always carefully attending to all manner of highly dubious conspiracy theories is unlikely to prove the most effective way to make accurate and timely decisions about public safety. Rather, public policy decisions should only be based on 'live' possibilities, where these are possibilities that are sufficiently probable. The correct threshold of probability will be that which, relative to the stakes and other factors, has the greatest expected lexicographic value when employed in decision-making. This, in turn, will be a matter of balancing contingent factors about the stakes at play in such decision contexts, the quality of the information available, the abilities of the decision-maker to accurately weigh evidence and moral considerations, and the time and resource pressures surrounding the decision. Once this threshold is set, the official will adopt a probabilistic threshold that, when followed as a rule, leads her to maximise the expected moral value of public safety, with the expected moral value of passenger satisfaction acting as a tie-breaker between equally choice-worthy threshold values.[18]

Note that lexical priority theories can adopt this approach because they are not, strictly speaking, Option Absolutist moral theories. Such theories categorically hold that some kinds of actions are always morally prohibited, irrespective of what other options are available to the decision-maker.[19] Since prohibition does not seem to come in degrees or to be amenable to aggregation, Option Absolutists cannot optimise in the face of risk. By contrast, lexical priority theories allow moral considerations of the same ranking to be traded-off against each other, making room for decision-makers to adopt a probabilistic threshold to optimise their conformity to lexically prior considerations. Under all but the most rarefied circumstances (those involving theoretically ideal agents in ideal situations), a context-variant, probabilistic threshold will best allow decision-makers to navigate moral decision-making under risk.

In what follows, we will equip lexical priority theories with a context-variant, odds-based threshold for determining which states to include in deliberations and which to exclude.[20] Where standard probability threshold approaches assess whether

here is silent on this point, and instead offers a pragmatic justification based on our bounded cognitive capacities and the importance of efficient versus accurate decision-making. See also: Jacob Ross and Mark Schroeder, "Belief, Credence, and Pragmatic Encroachment," *Philosophy and Phenomenological ...* 88, no. 2 (2012): 259–288. Secondly, Smith's account merely permits, but does not require, ignoring probabilities below the threshold. The justification offered in this chapter is that we are morally required to ignore sufficiently improbable possibilities, since doing so best ensures that we conform to our moral requirements.

18. One might worry that this approach to fixing the threshold leads to an infinite regress of decisions about how to decide. There are resources available to avoid this problem, notably: Hanti Lin, "On the Regress Problem of Deciding How to Decide," *Synthese* 191, no. 4 (2014): 661–670.

19. Jackson and Smith, "Absolutist Moral Theories," p. 268. See Chapter 1.

20. Hanti Lin and Kevin T. Kelly, "Propositional Reasoning that Tracks Probabilistic Reasoning," *Journal of Philosophical Logic* 41, no. 6 (2012): 957–981; Hanti Lin and Kevin T. Kelly, "A geo-logical solution to the lottery paradox, with applications to conditional logic," *Synthese* 186

or not a given state is sufficiently probable, the odds-based threshold evaluates whether a state is sufficiently more probable than its most probable alternative. This latter approach ensures not only that a consistent set of states is included in the decision problem, but also that the probabilities of these states are updated in accordance with the Bayesian rules of belief revision. Using this threshold, the official will include in her deliberations the states that are not sufficiently less probable than any other states. She will exclude from her deliberations any state that is sufficiently less probable than some other state. She will then update her beliefs about the included states, normalising the probabilities according to the rules of belief revision, and then maximise expected moral value.

To illustrate, let the relevant odds-based threshold, t, equal $1:99$. In the case of the Low Risk Problem, the probability that the public will be careless if she raises the speed limit is less than 0.01. As such, it is insufficiently probable. Given this information, the official should exclude the possibility that the public will be careless if she increases the speed limit, update her credences, and frame her decision problem as follows (Table 2.4).

Action	*Careless*	*Careful*	*Expected Value*
Increase speed limit	$-1000, 1$ $p = 0$	$0, 1$ $(1 - p) = 1$	$0, 1$
Maintain speed limit	$-2, -1$ $q = 0.45$	$0, -1$ $(1 - q) = 0.55$	$-0.9, -1$

Table 2.4: The Low Risk Problem Avoided

Even though increasing the speed limit has a positive probability of violating a lexical priority, this fact is excluded from the official's deliberations. Updating her beliefs on the fact that the public is sufficiently unlikely to be careless (given that she increases the speed limit), the official maximises expected moral value within the revised decision problem by increasing the speed limit.[21] In this way, the official upholds her commitment to lexical priorities while also avoiding the Low Risk Problem. However, as we shall see, positing any kind of probabilistic threshold seems to raise a further problem for lexical priority theories.

(2012): 531–575.

21. This approach is similar to that proposed in Hawley, "Moral Absolutism Defended." One main difference is that Hawley's threshold approach, being based on a fixed probability value rather than an odds-based threshold, does not obey plausible principles of belief revision and also encounters Lottery Paradox-style Agglomeration Problems. See: Section 2.5.

2.5 The Agglomeration Problem

The Agglomeration Problem consists of cases where a lexical priority theory offers inconsistent verdicts about what ought to be done, depending on whether moral considerations are responded to separately or together. There are two versions of the Agglomeration Problem: one conjunctive and the other disjunctive. As critics of lexical priority theories note, both versions of the problem result from positing a probabilistic threshold. They conclude that lexical priority theories therefore cannot appeal to probabilistic thresholds to avoid the Low Risk Problem. Below, I will explain how lexical priority theories can avoid both versions of the Agglomeration Problem, without abandoning a probabilistic threshold.[22]

The Conjunctive Agglomeration Problem involves cases where each of an agent's options has some moral property (say, that of respecting a lexical priority), but the conjunction of those options has a different property (say, that of violating a lexical priority). To illustrate, suppose that the official has just received additional information about the carefulness of the public, indexed to weekday behaviour and weekend behaviour. To simplify the following discussion, we will assume that the official should raise the speed limit if and only if there is a live possibility that the public is careful all week; otherwise, maintaining the speed limit is the most appropriate option available.

Given this simplifying assumption, if the official were to apply a probabilistic threshold, t, she might face the following dilemma. Suppose that the probability that the public will be careful on weekdays is greater than or equal to t (hence, it is a live possibility) and the probability that the public will be careful on weekends is greater than or equal to t (hence, a live possibility). It therefore seems that the official should increase the speed limit. However, it also seems possible that these states are incompatible (or, at least, anti-correlated), such that the probability that the public will be careful on both weekdays and weekends could be less than t (and, hence, not a live possibility), in which case she should not be open to increasing the speed limit. What ought she to do? Lexical priority theories appear to offer no guidance about what should be done. As it stands, the official seems to face an especially implausible kind of moral dilemma: if she decides on the basis of the time periods taken separately, she should perform an action that would be ruled out if she were to decide on the basis of the time periods taken together (Table 2.5).[23]

In response, it is first worth noting that nothing about this problem relies on the notion of lexical priority. The problem is more generic: when a threshold is posited,

22. This discussion addresses synchronic versions of the Agglomeration Problem. For a discussion of strategies for solving the diachronic versions of the Agglomeration Problem, see: Chapter 3.

23. As noted in: Jackson and Smith, "Absolutist Moral Theories," pp. 276-278; Huemer, "Lexical Priority and the Problem of Risk," pp. 336-339.

State	Probability (p, q, r, given **Increase Speed Limit**)
Public is careful on weekdays	$p \geq t$
Public is careful on weekends	$q \geq t$
Public is always careful	$r < t$

Table 2.5: The Conjunctive Agglomeration Problem.

there can be cases where the relevant conjuncts fall on one side of the threshold, but the conjunction falls on the other.[24] This should be encouraging for lexical priority theorists: the Agglomeration Problem is, strictly speaking, orthogonal to the concept of lexical priority; the problem's source – as well as its solution – lies elsewhere.

Note, also, that unlike the previous problems, the Conjunctive Agglomeration Problem is based on an informal description of the decision situation. This is important because, as it turns out, the problem disappears when it is placed in a more formal decision-theoretic framework. The informality of the Conjunctive Agglomeration Problem exploits an ambiguity in how to describe the relevant states in a decision problem. The decision-theoretic model forces a resolution of this ambiguity, thereby preventing the dilemma from arising.

In this case, the ambiguity driving the Conjunctive Agglomeration Problem is that there are two ways for the public to be careful: by being careful during a particular time period only (say, weekdays) or by being careful during that period and the remaining time period (say, weekdays and weekends). Since different degrees and types of risk may be associated with these different possibilities, they should be distinguished. A more accurate specification of the states of the decision problem would be as shown in Table 2.6.

State	Probability (p, q, r, s, given **Increase Speed Limit**)
Public is always careful	$p < t$
Public is careful on weekdays only	$q \geq t$
Public is careful on weekends only	$r \geq t$
Public is always careless	$s = 1 - p - q - r$

Table 2.6: A more accurate specification of the states.

Applying the relevant probabilistic threshold truncates the decision problem to rule out states whose probability is less than t. This rules out the possibility that the public is careful all week. For simplicity, we will also assume that the probability that the public is careless all week is also less than t. Having ruled out these possibilities

24. See, for example, Henry Kyburg's epistemological puzzle, the Lottery Paradox, which is structurally almost identical to the Agglomeration Problem. See: H.E. Kyburg, *Probability and the Logic of Rational Belief* (Middletown, CT: Wesleyan University Press, 1961).

from her deliberations, the official updates her beliefs and maximises expected moral value (Table 2.7).[25]

Action	*Public is careful on weekdays only*	*Public is careful on weekends only*	*Expected Value*
Increase speed limit	$-10, 1$ p	$0, 1$ $(1-p)$	$-10p, 1$
Maintain speed limit	$0, -1$ q	$0, -1$ $(1-q)$	$-0, -1$

Table 2.7: The Conjunctive Agglomeration Problem Avoided.

Given the arbitrary values above, the official should, in this case, maintain the speed limit. The more important point, however, is that there is no dilemma.

Moreover, there will be no dilemma, whatever values are plugged into the decision problem. By disambiguating the informal description of the case, we created a well-defined decision problem relative to which the official can maximise expected moral value. The Conjunctive Agglomeration Problem has been solved.

The Disjunctive Agglomeration Problem relies on a similar ambiguity about the relevant state-space of a decision problem. To illustrate, let us again assume that the official should raise the speed limit if and only if there is a live possibility that the public is careful all week; otherwise, maintaining the speed limit is the most appropriate option available.

Suppose now that the official's evidence indicates that it is sufficiently improbable that the public is careless on weekdays, such that she should exclude that possibility and be provisionally open to raising the speed limit. If it is also sufficiently improbable that the public is careless on weekdays, then it seems that she should therefore hold that there is a live possibility that the public is always careful and so should raise the speed limit.

Given this evidence, however, the following situation is also possible: it may also be sufficiently probable that the public is careless during weekdays or the weekend. Even though the probability of each disjunct falls below the threshold, the sum of their probabilities may be above it. In this case, the official cannot exclude the possibility that the public is careless during at least one of the time periods, whichever it happens to be. Given these assumptions, it appears that the official faces a strange kind of moral dilemma: if she decides on a 'time-period-by-time-period' basis, she ought to raise the speed limit; if she decides 'all together', then

25. As mentioned in the previous section, using an odds-based threshold rule, the official's belief revision will obey Bayesian Conditionalisation (meaning that the probabilities are updated in proportion to their prior relative probabilities, renormalised so as to sum to 1. See: Lin and Kelly, "A geo-logical solution to the lottery paradox, with applications to conditional logic"; Lin and Kelly, "Propositional Reasoning that Tracks Probabilistic Reasoning"

she ought to maintain the speed limit. It seems that whatever she chooses, she will be acting wrongfully.

As we shall see, the Disjunctive Agglomeration Problem exploits the fact that the concept of lexical priority underdetermines the appropriate decision-theoretic representation of a moral decision problem. The Disjunctive Agglomeration Problem consists of cases where there are conflicting but seemingly equally eligible ways of specifying the relevant states of a decision problem. For example, taken on a 'time-period-by-time-period' basis, the relevant states (S) of the official's decision problem are:

S1: The public is always careful.

S2: The public is careful on weekdays but careless on weekends.

S3: The public is careless on weekdays but careful on weekends.

S4: The public is always careless.

This specification of the decision problem suggests that considerations of public safety pertain to particular, identified, risks. For instance, it may be that different subgroups or individuals in the population are exposed to heightened risk at different times periods (such as children on weekends), and the official may be particularly concerned to ensure that disproportionate risk is not imposed on those subgroups or individuals. That the public is careless at some time or other is not relevant, given this specification of the decision problem.

However, having specified the decision problem as above, standard rehearsals of the Disjunctive Agglomeration Problem then make salient an alternative framing of the decision problem, whereby unidentified risk is morally relevant. This leads to a different specification of the decision problem's relevant possibilities, such as:

S1*: The public is always careful.

S2*: The public is at least sometimes careless.

On this approach, evidence about whether the public is careful is only relevant insofar as it indicates whether S1* or S2* is true. The official need not bother about which particular time period the public is careless during, so long as she is sufficiently certain that there is some time during which the public faces an increased risk of death.

The Disjunctive Agglomeration Problem thus raises a valuable question about what lexical priority theories, and moral theories in general, really care about. In the official's case, the question is: should she care about identified or unidentified

risk imposition? Strictly speaking, the concept of lexical priority is silent about this. The Disjunctive Agglomeration Problem thus shows that lexical priority theories are under-theorised in this important respect. In order to provide determinate advice in risky situations, such theories must explain how we should frame decision problems. To do this, they must spell out their substantive commitments in more detail.

As it happens, when it comes to the ethics of distributing risk, it is usually not enough to know that some quantum of risk is being imposed, whoever it might befall; it is often important to know whether particular groups or individuals are bearing an unfair burden of the risk. Such fairness considerations would support a framing of the decision problem in terms of identified risks. On the other hand, in cases where there is a diffuse risk that affects the population equally, an official may be concerned solely with the probability that the risk will eventuate somehow or other (this might apply, for example, when evaluating the risk of a catastrophic nuclear disaster). The framing of the decision problem is determined, in the end, by the details of the background moral theory.

Once the framing is settled, however, the Disjunctive Agglomeration Problem does not arise: the odds-based threshold ensures that only a consistent set of possibilities is included in the decision problem.[26] Once the decision problem is specified, the official should maximise expected lexicographic moral value.

2.6 Further Uncertainties

We have seen that lexical priority theories have the resources to avoid the various decision-theoretic objections that have been raised against them. The Permissiveness Problem relied on the mathematical oddities of infinite values, but lexical priorities can be modelled without infinite values. The Low Risk Problem was driven by the assumption that decision-makers should always attend to all possibilities, no matter how improbable. Lexical priority theories can explain why some possibilities should not be factored into a decision problem. The Agglomeration Problem exploited the fact that the concept of lexical priority underdetermines the appropriate framing of decision problems. Once lexical priority theories decide on how a given decision problem should be framed, they avoid the Agglomeration Problem. Each of these problems was rectified without having to retreat from the idea that some kinds of moral considerations cannot be defeated by any number of particular other considerations. In each case, the problem was actually a symptom of the chosen decision-theoretic representation. There are a few ways of interpreting this result.

One interpretation takes the failure of any such decision-theoretic critique as a

26. This is one of the key results in Lin and Kelly, "Propositional Reasoning that Tracks Probabilistic Reasoning"; Lin and Kelly, "Propositional Reasoning that Tracks Probabilistic Reasoning"

foregone conclusion. After all, it seems that we know (indeed, *a priori*) that the truth of moral theories depends on the adequacy of their substantive justifications, not the adequacy of their decision-theoretic representations. Faced with a problematic decision-theoretic representation of a moral theory, it is open to the moral theorist to say: 'So much the worse for your model!' On this view, the decision-theoretic critique of lexical priority theories was bound to fail because it does not engage with the substantive justifications of the lexical priorities in question.

However, even if the above is true, this does not mean that decision-theoretic critiques of moral theories are without value. In the process of responding to these objections, lexical priority theories have been forced to clarify their substantive commitments: do they really consider some considerations to be infinitely more valuable than others? Do they require that we always attend to arbitrarily small probabilities in our decision-making? Do they care about identified risks (such as protecting the lives of particular people) or unidentified risks (protecting the lives of people in general)? Even if the decision-theoretic critique fails, as argued, it has nevertheless succeeded in revealing these under-theorised aspects of lexical priority theories.

Indeed, once we scratch below the surface of lexical priority theories, we discover further puzzling features, depending on the particular kind of substantive justification at play. For example, consider the view that lexical priorities outweigh other considerations. How is this to be modelled? We have seen that positing infinite values is highly problematic. Instead, lexical priority theories could perhaps adopt a different kind of value function. One idea is to hold that the marginal moral value of upholding a lesser consideration diminishes asymptotically towards a limit, such that upholding any number of lesser considerations never has as much moral value as upholding a higher consideration.[27] This kind of moral value function could ensure that, in the official's case, no matter how many millions of passengers may benefit, pleasing an additional passenger will never outweigh the moral importance of a person's life.[28] The question is: why should the moral importance of an additional passenger's pleasure diminish due to external factors, like the number of other passengers who happen to also be benefiting from slightly more exhilarating rides? If moral value is conditional in this way, what exactly are the conditions? These questions become salient only once we try to understand how lexical priorities

27. For discussion of this idea, see: Erik Carlson, "Organic Unities, Non-Trade-Off, and the Additivity of Intrinsic Value," *Journal of Ethics* 5 (2001): 335–360. As applied to ethics and risk, see: Chapter 3.

28. This idea has been explored with respect to 'the good' by: Erik Carlson, "Aggregating Harms — Should We Kill to Avoid Headaches?," *Theoria* 66, no. 3 (2000): 246–255; John Broome, "No Argument against the Continuity of Value: Reply to Dorsey," *Utilitas* 22, no. 04 (November 2010): 494–496. For an approach that applies to multiple kinds of values, see: Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, chap. 10.

operate in a decision-theoretic framework.

Or consider, instead, the cancellation approach to justifying lexical priorities. It holds that there is no moral value to acting upon some considerations unless we also conform to particular other considerations. For example, we might say that there is no moral value in the official deviating from her role as a protector of public safety to satisfy the passengers' need for speed: optimising passenger satisfaction has no objective moral value when it involves disregarding the lives of those who may suffer the consequences.[29] This approach has its own mysteries: for instance, how do we weigh considerations when we are uncertain if they have been cancelled? Is there such thing as partial cancellation of moral value?[30]

Finally, consider an exclusionary approach to justifying lexical priorities. This approach is silent with respect to whether lexical priorities outweigh or cancel other considerations; it instead argues that lexical priorities deliberatively *exclude* them.[31] That is, there are cases where we should act first and foremost on the balance of only some types of considerations (such as public safety), irrespective of how many other considerations of a particular kind (such as additional passenger satisfaction) are at stake. As with the cancellation approach, it is not at all obvious how exclusionary reasons operate in cases of uncertainty. Are they nothing more than useful guides for decision-making (similar to the threshold approach presented earlier)? Or, are they better understood as epistemic considerations that give us reason to change our credences that the world is one way or another? In any case, does an exclusionary reason's importance diminish with probability and, if so, can such 'diminished' reasons be weighed against the reasons they purport to exclude?

These are almost completely unexplored issues in moral theory. The reason is that they only become salient once we adopt a decision-theoretic perspective of lexical priority views. Although the recent decision-theoretic critique of lexical priorities fails, it has succeeded in revealing new, potentially important lines of inquiry in our moral theorising.

29. A related notion of conditionality of moral worth can be found in, for example: Immanuel Kant, *Groundwork for the Metaphysics of Morals*, ed. Allen W. Wood (New Haven; London: Yale University Press, 2002), sec. 1.

30. As far as I am aware, the closest work that systematically addresses this question is: John Horty, *Reasons as Defaults* (Oxford University Press, 2012). However, Horty's approach does not seem to be sufficiently general for the purposes of most lexical priority theories, since it avoids both probabilistic uncertainty and the idea that moral considerations can be meaningfully weighed against each other.

31. See: Joseph Raz, *Practical Reason and Norms* (Oxford: Oxford University Press, 1999). Thanks to Andrew Williams for discussion.

2.7 Conclusion

The decision-theoretic critique of lexical priority theories suggests that they have no acceptable moral decision rule for cases of uncertainty, and should therefore be abandoned. However, as shown, this critique relies on contentious modelling assumptions that lexical priority theories need not – and, indeed, should not – accept. By using additional resources in decision theory and spelling out their substantive moral commitments in more detail, lexical priority theories can guide us through risky situations.

Chapter 3

Priorities and Uncertainties

Introduction

In Chapter 2, we saw how lexicographic decision theory allows lexical priority theories to avoid the Problem of Risk. In this chapter (based on a co-authored paper with Seth Lazar), we discuss how orthodox expected value theory can also do so.[1] A benefit of this latter model is that it allows us to explore, with more precision, the broader spectrum of lexical priority relations that moral theorists might adopt.

3.1 Varieties of Priorities

Some think that in any choice between saving a life and averting headaches, you should always save the life, no matter how many people whose temporary, minor pain you could otherwise alleviate.[2] They are the heirs of a more extreme form of lexical priority, for which some moral considerations matter so much that they could never be forsaken, regardless of what else was at stake: let justice be done though the heavens fall.[3]

This chapter is about various kinds of lexical priority theories – all of which are deemed to be absolutist by their critics and therefore subject to the Problem of Risk. As in Chapter 2, we will define this genus to be any theory that holds that there are some higher and lower considerations, such that no amount of the lower

1. Lazar and Lee-Stronach, "Axiological Absolutism and Risk."

2. Dale Dorsey, "Headaches, Lives and Value," *Utilitas* 21, no. 01 (2009): 36–58; Frances M. Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (Cary, NC: Oxford University Press, 2006); Michael Ridge, "How to Avoid Being Driven to Consequentialism: A Comment on Norcross," *Philosophy & Public Affairs* 27, no. 1 (1998): 50–58; Scanlon, *What We Owe to Each Other*; Larry S. Temkin, "A Continuum Argument for Intransitivity," *Philosophy and Public Affairs* 25, no. 3 (1996): 175–210; Voorhoeve, "How Should We Aggregate Competing Claims?"

3. G. E. M. Anscombe, "War and Murder," in *Moral Problems: A Collection of Philosophical Essays*, ed. James Rachels (New York: Harper & Row, 1979), 393–407; John Finnis, *Moral Absolutes* (Catholic University of America Press, 1991); Gewirth, "Are There Any Absolute Rights?"

consideration can defeat some amount of the higher consideration. However, in this Chapter, we will identify various stronger and weaker versions of lexical priority. We argue that each of these versions can adopt orthodox decision theory to avoid the Problem of Risk.[4]

That said, however, we do not provide a complete justification of the assumptions underlying our decision-theoretic model. Nevertheless, we take comfort in the fact that these assumptions rest on fundamental questions of moral theory, for which no decisive answers currently exist.

3.2 Preliminaries

Following Chapter 2, we will assume that lexical priority theories can be Relational Absolutist theories (in the sense defined by Chapter 1). We will also speak of 'considerations' as a placeholder for reasons, values, or anything else that bears on choice in a similar way (our goal is to preserve some neutrality between these alternatives).[5] We will also talk about rankings of considerations in terms of the defeat relation: one consideration c_1 defeats another c_2 just in case, in a choice between them where all else is equal, you morally ought to serve c_1.

To capture the various shades of lexical priority, we will need to introduce various logical quantifiers. Let $\mathbf{C}$ be the set of higher considerations and C be its extension (that is, the particular considerations that belong in that set). Let $\mathbf{C}(T)$ be the set of *types* of higher considerations. Let $\mathbf{c}$ be the set of lower considerations and c its extension.[6] Let $\mathbf{c}(T)$ be the set of types of lower considerations.

To make matters more concrete, we will understand c to be the consideration of averting a minor, temporary headache for a single person (belonging to the type: trivial pain). C denotes the consideration of saving a single life (belonging to the type: preserving life). Often, our actions involve multiple considerations of the same or differing kinds. We refer to considerations that belong in the same set as 'amounts'. The phrase 'any number of c' means 'any positive number of individuals

4. In particular, we will draw on: Gustaf Arrhenius, "Superiority in Value," *Philosophical Studies* 123, nos. 1-2 (2005): 97–114; Gustaf Arrhenius and Wlodek Rabinowicz, "Millian Superiorities," *Utilitas* 17, no. 02 (2005): 127–146; Gustaf Arrhenius and Wlodek Rabinowicz, "Value Superiority," in *The Oxford Handbook of Value Theory* (Oxford University Press, 2015), 225–243; Carlson, "Aggregating Harms — Should We Kill to Avoid Headaches?"; Carlson, "Organic Unities, Non-Trade-Off, and the Additivity of Intrinsic Value"; Dale Dorsey, "Headaches, Lives and Value," *Utilitas* 21, no. 1 (2009): 36–58; James Griffin, *Well-Being: Its Meaning, Measurement and Moral Importance* (New York: Oxford University Press, 1986); Kartsen Klint Jensen, "Millian Superiorities and the Repugnant Conclusion," *Utilitas* 20, no. 3 (2008): 279–300.

5. For instance, we could cast the discussion in terms of 'claims': Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*; Scanlon, *What We Owe to Each Other*; Taurek, "Should the Numbers Count?"; Voorhoeve, "How Should We Aggregate Competing Claims?"; Lazar, "Limited Aggregation and Risk."

6. We are assuming, of course, that $\mathbf{C}$ and $\mathbf{c}$ are mutually exclusive.

such they they each receive only one headache'. The phrase 'any number of C' denotes 'any positive number of individuals such that they are not killed as a result of your actions'. Lastly, the phrase 'C has lexical priority over c' means that some number of C defeats any number of c. If C is 'saving a life' and c is 'averting a single minor temporary headache', then a theory posits a lexical priority when it says that for any number of tokens of c, you should serve C, saving the life rather than averting the headaches.

Lexical priority can vary in *scope* and *stringency*. Scope refers to the size of the set of types of considerations which the higher consideration defeats. Stringency refers to how many higher considerations are needed to instantiate the lexical priority relation. We refer to some combinations of these positions below:

Stringent-Wide Priority is maximally stringent and has maximal scope. It is maximally stringent because it holds that a single token of C defeats any number of c. It has maximal scope because it holds that the set of types of higher considerations is singleton and its priority ranges over the domain all other considerations. This position might say, for example, that one may never kill an innocent person, no matter the good that could thereby be achieved.

Stringent-Narrow Priority is maximally stringent, but has narrow scope. It holds that any number of C defeats any amount of c. However, it restricts $\mathbf{c}(T)$ to only trivial pains, and is silent about whether other types of considerations belong in that set. The 'life for headaches' case implies this kind of priority.

Weak-Wide Priority holds that some amount of C (greater than one) defeats any amount of all other considerations. For instance, we might hold that preserving some critical number of human lives on our planet (under relatively free and humane conditions, we might say) defeats any number of any other considerations.

Weak-Narrow Priority holds that some amount of C (greater than one) defeats any amount of some particular type of consideration.

We will argue that the above kinds of priority relations between considerations are consistent with orthodox decision theory. This will allow lexical priority theories a different way to avoiding the Problem of Risk.

Very roughly, orthodox decision theory tells us that, faced with imperfect information, we should identify our options and the possible outcomes to which each might lead, assign values to those outcomes, and probabilities conditional on taking that option, then multiply the two numbers together, before choosing the option for

which the sum of those products is greatest.[7]

A moral theory tells us the relative importance of the considerations at play in our choices. So long as its ranking of moral considerations satisfies particular structural conditions (such as transitivity), we can represent the importance of moral considerations by assigning a value of the outcomes in which they are upheld or violated. We can interpret this value as being merely representational (a convenient index for comparing considerations) or as a more metaphysically 'real' entity. Also, we could interpret value as being grounded on an individual's preferences or instead on some objective standard of importance. We will remain neutral on these metaphysical and metaethical points. We will simply illustrate that lexical priority theories plausibly satisfy the structural conditions of orthodox decision theory.

The first step is to model a lexical priority theory's moral value function for outcomes. Philosophers sometimes assume that the only way to model the value function of a lexical priority theory is to posit an infinite value difference between higher and lower considerations.[8]

Here, however, we will take an alternative approach. Rather than positing an infinite value difference between C and c, we will posit that the value of any amount of c cannot exceed some upper bound.[9] We thus deny the 'simple additivity' of c.[10] The most obvious alternative is that c has diminishing marginal value, which decreases asymptotically towards zero, so that the total value approaches a limit beneath the value of a single token of C. This reflects the fact that no amount of c can defeat some amount of C, while still holding that each c is morally important in its own right.

7. Decision theory is no more settled than moral theory. This is a simplified statement of the orthodox view, and it punts on difficult questions like whether evidential or causal decision theory is correct. For two overviews, see **Briggs2014**; Lara Buchak, "Decision Theory," in *The Oxford Handbook of Probability and Philosophy*, ed. Alan Hájek and Christopher Hitchcock (Oxford: Oxford University Press, 2016), 789–814.

8. This assumption underpins Norcross's early objection to absolutism (Alastair Norcross, "Comparing Harms: Headaches and Human Lives," *Philosophy and Public Affairs* 26, no. 2 (1997): 135–167; Alastair Norcross, "Great harms from small benefits grow: how death can be outweighed by headaches.," *Analysis*, no. April (1998): 152–158. It is also a background assumption in Huemer, "Lexical Priority and the Problem of Risk," and in parts of Isaacs, "Duty and Knowledge," Lord and Maguire, "The Implementation Problem for Deontology." For versions of this approach, see: Chapter 2.

9. We think the first to do so was Carlson (2000). But see also Arrhenius and Rabinowicz (2005a), Ken Binmore and Alex Voorhoeve, "Defending Transitivity against Zeno's Paradox," *Philosophy and Public Affairs* 76, no. 3 (2003): 272–279; Broome, "No Argument against the Continuity of Value: Reply to Dorsey"; Jensen, "Millian Superiorities and the Repugnant Conclusion"; Arrhenius and Rabinowicz, "Millian Superiorities."

10. By 'simple additivity', we mean the property whereby, holding the scale of value constant, every token of c adds as much to the total value as every other token of c.

3.3 Lexical Priority and Risk

This alternative is good news for lexical priority theories: bounded value functions are quite consistent with orthodox decision theory.[11] However, nobody yet has considered precisely how to implement this idea for moral decision-making under risk. That is our task in this section.

Suppose only two considerations, C and c, bear on the moral status of some choice, and that some amount of C outweighs any amount of c. You face a risky choice, in which you could realise some amount of C and some of c. To apply this Weak-Narrow Priority to decision-making under risk, we can do the same thing as any other orthodox decision theorist: consider the possible outcomes and assign them a moral value (MV), based on our objective moral theory.[12]

Say that the moral value of C increases linearly without upper bound, each token adding 20 more units of moral value. By contrast, c diminishes at the margin, approaching an upper bound of 99 units of moral value. So five tokens of C outweigh any amount of c, but some amount of c can outweigh four or fewer tokens of C. So, C has weak and narrow priority over c.

Now consider this decision problem, illustrated in Table 3.1, in which ϕ and ψ are actions, and A and B are possible states of the world.

	A	B
ϕ	$10C; 0c$ (MV=200)	0C;0c (MV=0)
ψ	$0C; 1,000,000c$ (MV≈99)	$0C; 1,000,000c$ (MV≈99)

Table 3.1: Weak-Narrow Priority and Risk

Suppose that $p(A|\phi)$ and $p(B|\phi)$ are both 0.5, as are $p(A|\psi)$ and $p(B|\psi)$. Then the expected moral value of ϕ-ing is 100, while that of ψ-ing is very close to 99. So you ought to ϕ, despite the astronomically larger number of tokens of c that you could realise by ψ-ing. The weak priority of C over c, however, does not entail that you will tolerate any risk to c to avoid any risk to C. Suppose the probability of B given ϕ or ψ is 0.51, while that of A is 0.49. Now you should ψ, because that has just short of 99 expected moral value, while the expected moral value of ϕ-ing is only 98. This basic approach extends to other kinds of lexical priority. We need only state the moral value function at stake, quantifying the priority relation, then input those values into our decision tables, to generate results consistent with standard decision

11. John W Pratt, "Risk Aversion in the Small and in the Large," chap. Chapter 19 in *Handbook of the Fundamentals of Financial Decision Making* (1964), 317–331; Kenneth J. Arrow, "The Theory of Risk Aversion," chap. 2 in *Aspects of the Theory of Risk Bearing* (Helsinki: Yrjo Jahnsonin Saatio, 1965)

12. This involves some contentious modelling choices; our aim is for a proof of concept, not to defend a specific set of numbers.

theory. Stringent priority relations are the same. Suppose that any amount of C defeats any amount of c. Each token of C has 100 moral value; the moral value of c cannot exceed a limit of 99. Now consider Table 3.2.

	A	B
ϕ	$2C; 0c$ (MV=200)	0C;0c (MV=0)
ψ	$0C; 1000000000000c$ (MV≈99)	$0C; 1000000000000c$ (MV≈99)

Table 3.2: Risking Weak Priorities

If the probabilities are again 0.5 across the board, then you should ϕ, despite the enormous number of tokens of c forgone. But if the probability of B given ϕ or ψ is 0.51, while that of A is 0.49, then the expected moral value of ψ-ing tops that of ϕ-ing ($\approx$ 99 to 98), and you ought to ψ. This, again, even though C has stringent priority over c. Stringent-Narrow Priority is structurally identical to Stringent-Wide Priority, except that where Stringent-Narrow Priority posits a specific consideration C that defeats a specific consideration c, Stringent-Wide Priority holds that a single token of C outweighs all other considerations. So what goes for Stringent-Narrow Priority also goes, *mutatis mutandis*, for Stringent-Wide Priority. This yields an interesting result. Stringent-Wide Priority starts to look much less extreme in risky cases. Even if C defeats all other considerations, there is no need to think that any prospect of C defeats all possible prospects of all other considerations. This might make Stringent-Wide Priority more palatable. Even if no number of all other considerations can outweigh a single token of C, they can do so when the latter is sufficiently unlikely.

This suggests that lexical priority theories face no special challenges when using orthodox decision theory. Of course, when we consider more complex decision problems, it would quickly become obvious that actually using this decision theory would be hellishly difficult. Although we take this concern seriously, it is a general problem. Moral philosophy alone cannot provide a decision procedure for realistic choices. But it can offer a criterion of subjective permissibility.[13] Lexical priority theories are no worse placed than others to do this.

3.4 Debunking the Objections

With this approach in hand, we can turn to the problems that critics have raised for applying lexical priority theories to risky choice situations. Swayed by a presupposition that such theories must posit infinite value differences, they have furnished

13. Lazar, "Deontological Decision Theory and Agent-Centered Options"; Lazar, "In Dubious Battle: Uncertainty and the Ethics of Killing."

them with a variety of different decision rules focusing on the idea of a probability threshold, in this way: suppose that C is your reason to abide by a constraint, and $p(\neg C)$ is the probability that ϕ-ing will breach the constraint. Critics have assumed that lexical priority theories must accept the following kind of decision rule: it is permissible to ϕ only if $p(\neg C)$ is less than some threshold. For example, it is permissible to risk killing an innocent person only if the probability that the person is innocent is less than p. These critics have then run a series of objections to highlight the problems with this kind of threshold view. Lexical priority theorists, for their part, have uniformly shared their assumption that orthodox decision theory cannot work for them, and have sought to defend some version or other of a threshold approach.[14] Our first and perhaps most important point is this: this dialectic has been labouring under a mistake. Lexical priority theories, like other theories, can use orthodox decision theory to extend their views to decision-making under risk. To be sure, there are many problems with using rational decision theory as a model for moral decision theory. But the inability to accommodate plausible lexical priorities is not one of them.

Our second key observation: the debate about lexical priorities and risk has been too much driven by Jackson and Smith's focus on Stringent-Wide Priority. We cannot think of a single influential contemporary deontologist who, in published work, endorses this view. Even Robert Nozick (who stretches the term 'contemporary) countenanced letting side-constraints be outweighed, if the stakes were high enough.[15] This misunderstanding is particularly important, since all of the most persuasive cases used by Jackson and Smith, and Huemer, involve high stakes whatever one does, often involving risking killing the innocent in order to save many more lives, or more realistic but structurally similar cases from war, or from criminal justice.[16] These examples are, to our minds, unhelpful for understanding the implications of risk for lexical priority theories, because only the most extreme absolutist would argue that we may never harm the innocent, no matter how great the good we can realise thereby, especially given that there is some probability that one's target is not in fact innocent.[17] What's more, as we have seen, even if one's objective moral theory was committed to Stringent-Wide Priority in cases like these, that would not entail that it would be equally rigorist for decision-making under risk.

Third: the most telling criticisms levelled by Huemer, and Jackson and Smith,

14. See also Bjorndahl, London, and Zollman, "Kantian Decision Making Under Uncertainty: Dignity, Price, and Consistency."

15. Nozick, *Anarchy, State, and Utopia*, p. 30

16. Jackson and Smith, "Absolutist Moral Theories"; Lord and Maguire, "The Implementation Problem for Deontology"; Huemer, "Lexical Priority and the Problem of Risk."

17. See Seth Lazar, "Risky Killing and the Ethics of War," *Ethics* 126, no. 1 (2015): 91–117.

have to do with problems of agglomeration that beset threshold-based views of subjective permissibility. Focusing on the kind of case introduced at the start of this section, these objections presuppose that absolutists must commit to a single value for $p(\neg C)$ (the probability that your action breaches a constraint), such that if it exceeds that threshold, the action is subjectively impermissible. The objection starts by claiming that this threshold must be arbitrary, then proceeds to consider cases in which one must perform multiple acts, each with some probability of breaching the constraint. Each of these assumptions can be questioned. Lexical priority theories need not endorse any contextually-invariant thresholds. For a given decision problem involving risking breaching a constraint C, if matters are simple enough then we can probably extract a threshold t such that if $p(\neg C)$ exceeds t, then it is subjectively impermissible to proceed. But this would be misleading; we could do the same for any maximising decision rule. According to orthodox decision theory, if utility is linear with monetary value, then placing a \$1 bet on a coin that will pay \$2 if it lands heads, nothing otherwise, is rational if and only if $p(heads) = 0.5$. But obviously we wouldn't describe 'maximise expected utility' as a threshold rule. Insofar as one insists on extracting a threshold from the lexical priority theory, it is no more arbitrary than this threshold extracted from rational decision theory. Where it sits depends on the stakes, as it obviously should. If the probability of heads is 0.499999, then the bet is irrational, regardless of how close you are to 0.5.

A further corollary: since the 'threshold' emerges from an expected moral value calculation given what is at stake, if we do extract one, it will depend on what is at stake in that particular decision problem. Several of the cases put forward by critics miss this point, positing a single value for the threshold regardless of whether one's action risks (for example) killing one innocent person or two. And if you are contemplating saving two groups of innocent people, rather than just one, then that might also affect the stakes, and so the implicit threshold at which it becomes permissible to proceed. The same will be true if more than one higher-order consideration is at stake. So all the examples used so far are ill-framed.

Ultimately we agree that these critics, albeit through a glass darkly, have identified a significant challenge. However, they have misdiagnosed the problem. The real problem does not derive from a commitment to thresholds that lexical priority theories need not make. Instead, it derives from a set of fundamental questions that go far beyond the notion of lexical priority, and must be addressed by any moral theory: namely, which units of action are the proper focus of moral enquiry: acts or campaigns? We concede, though, that this problem has particular purchase for those who reject the simple additivity of value. We come back to this in Section 5.

One objection to our view is that this an ersatz kind of lexical priority.[18] In

18. Huemer (2010) implies something like this.

response, it is worth making one feature of our approach clear. If the moral value of the lower consideration is bounded, and diminishes at the margin, then, as the risk of breaching the higher consideration increases, it will take more and more of the lower consideration to counterbalance that risk.

Imagine, for example, that not killing an innocent person defeats saving innocent lives, so that it would be impermissible to kill an innocent person no matter how many lives you could save thereby.[19] (Few people hold such an extreme view, but let that slide.) But suppose that the upper bound of the value of saving innocent lives is equivalent to nine-tenths of the moral value of knowingly killing an innocent person. So, if the probability that this action will kill an innocent person is 0.1, say, then there is some number of lives certainly saved that could make that action permissible. Note, though, that as the probability that this action will kill an innocent person increases, the number of people whom you would have to be sure of saving in order to permissibly run this risk increases by a disproportionate amount, as each additional life saved adds less to the action's overall moral value. Once you reach a 0.9 probability of killing an innocent person, no number of lives certainly saved would make it permissible to proceed. To repeat, we think this kind of view is deeply implausible. But it is clearly not ersatz.

3.5 Additivity and Individuating Options

Critics have argued that lexical priority theories are uniquely ill-equipped to address decision-making under risk. They are wrong. Such theories can use orthodox decision theory as well as anyone else, if they deny the simple additivity of lower considerations. However, this approach comes with its own problems.

Interestingly, though, the reason it does so is quite different from what its critics have assumed – and it really has nothing to do with risk. The basic problem is this. Some moral (and rational) decision theories include elements that generate divergent judgements of what seem to be morally identical phenomena. Specifically, they enjoin us to differently assess a sequence of acts, depending on whether we consider each act in isolation from the others, or the whole sequence as a campaign. In decision theory this is sometimes described as a conflict between 'local' and 'global' rationality. It is relatively common: philosophers have shown how vagueness, infinite utilities, rejection of the reflection principle, and risk-averse attitudes can all lead to clashes of this sort.[20] And in moral theory, the phenomenon is not confined to

19. See Lazar, "Risky Killing and the Ethics of War"; Lazar, "Limited Aggregation and Risk."

20. Frank Arntzenius, Adam Elga, and John Hawthorne, "Bayesianism, infinite decisions, and binding," *Mind* 113, no. 450 (2004): 251–283; Buchak, *Risk and Rationality*; Adam Elga, "Subjective Probabilities should be Sharp," *Philosophers' Imprint* 10, no. 5 (2010): 1–11; Brian Hedden, "Options and Diachronic Tragedy," *Philosophy and Phenomenological Research* 87, no. 1 (June

so-called absolutist moral theories. Indeed, Frank Jackson himself (with Robert Pargetter) has argued that their Professor Procrastinate is one such case.[21] Jackson thinks that Procrastinate ought to ACCEPT THE REQUEST TO REVIEW THE PAPER AND WRITE THE REPORT ON TIME. But given that he will not write the report on time if he accepts, he ought not to ACCEPT THE REQUEST TO REVIEW THE PAPER. If we consider the sequence of actions, then he acts impermissibly by refusing to referee the paper; if we consider the individual action, he acts permissibly. Many other such cases have been put forward that make the same point.

To convey the specific form of this problem for lexical priority theories, we need a cleaner case than has been offered so far by critics focused on debunking Stingent-Wide Priority, a view that nobody holds.[22]

In this case, C is letting an innocent person die, who would otherwise live a happy and long life; c is a unit of financial benefit to a person, equivalent to the value to an already wealthy person of \$100.[23] Each token of c is one (different) person benefited in this way. Plausibly, no amount of c could outweigh a single token of C. No matter how many well-off people we could benefit by \$100, if we could instead save an innocent person's life, then we should do so. This looks like a plausible kind of Weak-Narrow Priority.

Assuming that C is not infinitely superior to c, for some probability p and some number of tokens n, one must be indifferent between $p(C)$ and nc. To fix things, let's stipulate that the moral disvalue of one token of C is $-9,999$, while that of c diminishes at the margin, approaching 10 at the limit as n approaches infinity. On these numbers, it would be permissible to take a 0.001 risk of failing to save a life in order to give n people \$100, for a high enough value of n.

Now suppose that on Monday morning, if you press BUTTON 1, you can achieve nc for sure, but only by running a 0.0009 risk of killing an innocent person, A. Since the value of nc is greater than that of $0.0009C$, you proceed. Then on Monday afternoon, you face the same decision problem: pressing BUTTON 2 will realise nc for a different set of beneficiaries, but subject a different possible victim, B, to a 0.0009 risk of death. The decision problem being the same, you proceed again.

On Tuesday you face two identical decisions at the same time. The probabilities are the same as before. If you push BUTTON 1, you expose only A to a 0.0009

2013): 1–35.

21. Frank Jackson and Robert Pargetter, "Oughts, Options, and Actualism," *The Philosophical Review* 95, no. 2 (1986): 233–255; Holly S. Goldman, "Dated Rightness and Moral Imperfection," *The Philosophical Review* 85, no. 4 (1976): 449–487.

22. Nonetheless, our case is indebted to those cases, especially as advanced by Huemer, "Lexical Priority and the Problem of Risk"; Jackson and Smith, "Absolutist Moral Theories"; Alastair Norcross, "Two Dogmas of Deontology: Aggregation, Rights, and the Separateness of Persons," *Social Philosophy and Policy* 26, no. 01 (2009): 76.

23. That is, we stipulate that the \$100 won't make the different to the satisfaction of any important interests

risk of death, achieving nc (that is, benefiting n people by c amount). If you push Button 2, you expose only B to a 0.0009 risk of death, achieving nc. And if you push Both Buttons, you expose both A and B to a 0.0009 risk of death, achieving $2nc$. If you push Neither Button, A and B are not exposed to risk, and nobody gets $100. In this case, you would be permitted to push Button 1 or Button 2, but not both. If you were unable to push only one of the buttons, you would have to do nothing.

But now suppose that you chose to push Button 1 on Tuesday morning, but in the afternoon discover that Button 2 can be pressed. Then, considering only the difference you can now make to the world, it would be permissible to push Button 2, having pushed Button 1.

This result is supposed to be counterintuitive: how can it make a difference, the critic asks, whether one pushes Both Buttons simultaneously, or does one in the morning, and one in the afternoon? It seems that by simply repackaging or disaggregating our options, we can generate different verdicts.

In properly cleaned-up cases, these results may not be so counterintuitive. But suppose that they are. The task, then, is to diagnose what is causing these results, to ask whether this problem is unique to absolutism, and then to sketch a response. We will proceed in that order.

Critics think that this problem of divergent verdicts on individual acts and campaigns composed of just the same acts derives from a problem of agglomeration, itself grounded in lexical priority theories' need to invoke a threshold in their moral decision theory.[24] But we have shown that the problem arises even when we do not adopt a threshold approach to decision-making under risk. For our approach, the source of the problem is the rejection of simple additivity – and it has nothing to do with risk.

If c has a diminishing marginal moral value function, then the weight of any token of c depends on how many other tokens of c are also in play. Say that a single token of c, on its own, has X value. If the c moral value function is concave and increasing, then X is the most a token of c can ever weigh. If there are multiple tokens of c, then each additional token must weigh less than the last. Now, imagine a sequence of acts $\phi_1, \phi_2 ..., \phi_{10}$, in each of which a single token of c is at stake. The aggregate moral value of those ten instances of c will be $10X$. But if ten tokens of c are at stake in a single act, ψ, they must together weigh less than $10X$, because of the diminishing marginal moral value function. Suppose that both the sequence, $\phi_1 ..., \phi_{10}$ and the act ψ involve the same costs, which are equivalent to just less than $10X$. Then it would be permissible to carry out the sequence, but not the single act, despite their involving the same expected costs, and the same expected benefits.

24. See also Chapter 2.

And this seems counterintuitive: if the costs and benefits are what really matters, then why should we care whether they are realised serially or at once?[25]

One response: this is not news! It's a basic feature of endorsing any holistic value judgements. The objection simply thumps the table and insists on additivity. But if the case for rejecting additivity is strong, as examples like 'life for headaches' suggest it is, then perhaps we should tolerate these intuitive costs. None of the competing moral theories come for free: endorsing additivity means endorsing the counterintuitive verdict in life for headaches.

A second response: these objections raise a fundamental question in both moral theory and decision theory, as to what the proper units of evaluation are. For if we should judge the sequence of acts $(\phi_1..., \phi_{10})$ as a whole, or if we should disaggregate the compound act into its constituent parts, then these inconsistencies would not arise. As already noted, precisely the same question arises in rational decision theory.

Some defenders of lexical priorities have implicitly noticed this possibility, and advocated for either the 'disaggregating' or the 'bundling' approach. Much of the subsequent discussion has shown that neither of these seems to be uniformly successful. In what remains of this section, we will suggest a way forward for other 'absolutist' theories and everyone else, too, since everyone needs an account of when campaigns or acts are the proper object of moral evaluation.

At this stage, it seems that one size does not fit all: sometimes we should disaggregate; sometimes bundle.[26] The trick is to find a principle that distinguishes these cases.

Suppose that 100 people are under threat of a terrorist attack. There are ten suspected terrorists. For each individual, the probability that he is part of the attack is 0.9. The terrorists have built in redundancy: if any one of them lives, the 100 will surely die. In the first scenario, the apparent terrorists are collocated – you can kill

25. See also Norcross, "Two Dogmas of Deontology: Aggregation, Rights, and the Separateness of Persons."

26. This approach is inspired by discussion with Sergio Tenenbaum, whose approach is to argue that both levels of permissibility matter (simultaneously): see Tenenbaum, "Action, Deontology, and Risk: Against the Multiplicative Model." Douglas Portmore has also developed work in this vein, in which he has argued for 'maximalism', in our terms roughly the view that campaigns, not individual acts, are the proper object of moral evaluation: see Douglas Portmore, "Maximalism versus Omnism about Permissibility," *Pacific Philosophical Quarterly* 98, no. 2017 (2017): 427–452. Brown (2017) offers a distinct defence of maximalism from Portmore's: see Campbell Brown, "Maximalism and the Structure of Acts," *Noûs* 0 (2017): 1–20 Kirkpatrick [2017] takes a similar position to Tenenbaum, arguing that the language of permissibility is ambiguous over whether campaigns or acts are the proper object of moral evaluation, and concluding that both levels matter: see James R. Kirkpatrick, "Permissibility and the Aggregation of Risks," *Utilitas* 30, no. 1 (2017): 107–119. Interestingly, while moral philosophers seem to lean ether towards maximalism or to saying that both kinds of permissibility matter, decision theorists seem to heavily favour focusing on very narrowly described acts: see Elga, "Subjective Probabilities should be Sharp"; Hedden, "Options and Diachronic Tragedy." This is *prima facie* evidence in favour of our preferred stance, that one size does not fit all.

them all in one go. In the second, they are separate, and must be killed separately. What is the proper unit of evaluation here? The individual act of killing? Or the campaign to stop the terrorist attack?

Here it seems clear that the campaign is morally fundamental. If we were to kill the first target, but leave the others to carry out the attack, then we would run a 0.1 risk of killing an innocent person for no good at all. The good that justifies the risk is achieved only if we impose the same risk nine more times.[27] In this case, whatever your method of killing the targets – in one go or sequentially – the campaign as a whole is the proper object of moral assessment. Each individual act of killing can bring about the good only if you kill the other targets.

But now suppose that the 100 victims are in ten groups of ten, and that each of the presumed terrorists will attack a different group. Then the good that justifies each 0.1 risk of killing an innocent person is independent from the goods that justify running the other risks. Whether we proceed with a single strike or hit the presumed terrorists one after the other, it seems clear that we may consider each act (or portion of the compound act) independently.

The general idea is this: when a given risk of a bad outcome is causally sufficient to realise enough expected value to justify that risk, then that risk can be assessed separately from whatever else is at stake. We can weigh the expected moral disvalue of running that risk against the expected value thereby realised. Suppose killing target A is causally sufficient to realise v^- expected moral disvalue, and v^+ expected moral value. If v^+ is greater than v^-, then killing A can in general be assessed in isolation from any subsequent acts. But if killing A is causally sufficient to realise v^+, then that expected value cannot count in favour of any subsequent risks, for fear of double-counting. When considering killing target B, we must set v^- and v^+ aside.

Now suppose that killing A is not causally sufficient to realise v^+ expected moral value (either it realises no expected value on its own, or it realises too little to justify v^-). In this case, if killing A is permissible it must be because it is necessary to some broader sequence of actions that does realise enough expected moral value to justify the expected moral disvalue of all the risks involved.

How does this (roughly sketched) approach fare with the standard counterex-

27. This case is similar to Huemer's example, in which we must choose between a 0.11 risk of killing one innocent person and a 0.1 risk of killing 1000 innocent people. He thinks that individualists such as Aboodi et al. will separate out the risks to each of the 1000, and so think it preferable to subject each individual to a 0.1 risk of innocent death than to subject one person to a 0.11 risk of innocent death. Set aside for now the fact that in one case the risk is that the party is innocent (death is assured), while in the other case we know the 1000 are innocent, but death is uncertain (on this difference, see Lazar, "Risky Killing and the Ethics of War"). In this case the risks to the 1000 are obviously not independent: you don't get any of the good unless you subject all 1000 to that risk. So obviously the expected disvalue will be much greater than running a 0.11 risk of killing one innocent person.

amples? Huemer and Jackson and Smith argue that if absolutists think that killing the innocent is absolutely prohibited, then they must be pacifists, and that if they think it is absolutely prohibited to punish the innocent, then they could not defend a workable criminal justice system. It bears repeating that these are not plausible examples. The goods achieved by just wars are weighty enough to justify risking killing innocent people; something similar is true for punishment. But forget that point. When thinking about wars, or criminal punishment, should we evaluate the practice as a whole? Or the individual acts of which it is composed?

Criminal punishment seems more like our second example. The potential benefits achieved by running a particular risk of punishing the innocent are independent of the potential benefits achieved by running further risks in other cases. So it seems right to judge individual acts of punishment in isolation: they are causally sufficient to realise enough expected moral value to justify the expected moral disvalue that they involve.

Wars look more like the first case. Although the permissibility of some acts in war is 'free-standing', in general fighting a war is permissible only if the many risks involved are justified by the overarching good for which they are necessary and sufficient. At the very least, we should also evaluate wars as whole campaigns. Many acts in war would not be justified if we focused only on the risks and benefits for which they are causally sufficient.

This is just a first step in a solution.[28] We are confident in the general prescription: everyone needs to think about when we should assess closely-connected acts together, and when we should disaggregate compound acts into their constituent parts. Our specific proposal is not exhaustive. It says simply that when a risky act is causally sufficient to realise some expected benefit, it can be considered in isolation from the campaign of which it is part, and assessed as permissible or impermissible. When one risky act depends on others to realise its expected good, then they must be assessed together.

But lexical priority theorists should take heart. First, everybody needs, and nobody yet has, an adequate account of whether acts or sequences of acts are the proper object of moral evaluation. So they have many companions in guilt. Second, non-additive, and otherwise holistic, value functions are far from outlandish. Indeed, the more controversial assumption might be that moral value is straightforwardly additive. Rational decision theory may have to presuppose that individuals have bounded value functions in order to address problems like the St Petersburg Game.[29] Most people already think that money has diminishing marginal value. Holistic and

28. A particularly pressing problem, for which we have no solution: what should we do when we are uncertain whether our act should be assessed as part of a campaign, or on its own?

29. Jeffrey, *The Logic of Decision*, p. 154; Nover and Hájek, "Vexing Expectations," 247-248.

contextual interactions between values are well-established.[30] One might even argue that presupposing additivity of value involves reifying and fetishising the mathematical notation with which we are representing our moral theories.

3.6 Conclusion

We have argued that critics of lexical priority theories have lumbered it with problematic moral decision rules. We have argued that these theories can adopt the rule of maximising expected moral value, without compromising their priorities.

In particular, we noted that the central objection to lexical priorities rests on a misdiagnosis. The problem is not with the use of thresholds, but instead derives from rejecting the simple additivity of value. Committing to a non-additive value function leads to theoretical complexities, but many others are in the same boat. The biggest problem is shared by everyone: are acts or sequences of acts the proper object of moral evaluation? Many non-standard versions of rational decision theory generate just the same kind of problem.

We have sketched an answer to that question. However, addressing it at length is a matter for further research. Our main aim has been to show that, for moral decision theorists who want to learn from decision theory, lexical priorities do not introduce any special new problems. Of course, some who endorse lexical priorities are dispositionally opposed to the whole value-theoretic project. And yet every moral theory has to have some way to cater for imperfect information. Those who reject decision theory must propose a more plausible alternative.

Of course, there are objections to our proposal. To use decision theory as we propose, one has to show that diminishing marginal value functions are plausible representations of the importance of moral considerations. Some think that there are general reasons to believe they cannot do so.[31] However, these objections must be careful not to make accuracy the enemy of usefulness: just as one cannot reject a useful map because contour lines do not literally exist on the landscape, one cannot object to our model on the basis of features that – though artificial – help to provide moral guidance in cases of uncertainty.

30. Jonathan Dancy, *Moral Reasons* (Oxford: Blackwell, 1993), xiii, 274 p.; Shelly Kagan, "The Additive Fallacy," *Ethics* 99, no. 1 (1988): 5.

31. See, for example, John Broome's discussion of 'strong separability', which implies that the wellbeing of multiple individuals cannot have diminishing marginal value, because the value of each individual's well-being is independent of the value of anyone else's wellbeing: see Broome, *Weighing Goods*. As a counterpoint, consider Kamm (2007), Scanlon (1998), Temkin (2012) and Voorhoeve (2014), who all argue that the moral significance of a person's interests in a choice can depend precisely on which other interests are at stake: see Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*; Scanlon, *What We Owe to Each Other*; Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*; Voorhoeve, "How Should We Aggregate Competing Claims?"

Chapter 4

Duty and Ignorance

4.1 Introduction

Holly Smith (2014) argues that deontological moral theories cannot correctly determine whether we ought to gather more information before acting.[1] For brevity, I will call this: the Problem of Ignorance. To illustrate, consider the following case:[2]

Diagnostics: You are a doctor in the emergency ward, faced with a patient who has a life-threatening condition. Your primary duty is to cure your patient. You also have a duty, albeit a less weighty one, to reduce her pain. You have the following choice:

Administer Drug: You can give the patient Drug A, which is a potential cure, or Drug B, which is merely a pain-reliever. If the patient is A-Receptive, then it will both cure her and relieve her pain. If the patient is not A-Receptive, then it will increase her pain and not cure her. In either case, Drug B will not cure the patient, but it will provide pain relief.

Alternatively, you can first:

Test, then Administer Drug: You can run a diagnostic test that will correctly determine whether the patient is receptive to Drug A (giving an output: A-Positive) or whether she is not so (giving an output: A-Negative). Then you can **Administer Drug**.

1. Smith, "The Subjective Moral Duty to Inform Oneself before Acting."

2. This case is intended to be structurally identical to Smith's lead case, which involves determining who a manager should lay-off. Here, I am instead using a medical case, similar to Jackson (1991), because its moral dimensions are arguably clearer. See: Jackson, "Decision-Theoretic Consequentialism."

> Running the test will have no cost (moral or otherwise): it will not cause the patient undue harm and will not jeopardise her or anyone's safety. As it stands, you do not have any beliefs about whether the patient is A-Receptive.[3]

We will assume that you morally ought to run the test: accepting free and reliable information that allows you to better treat your patient is something that any adequate moral theory should require. The challenge for subjective deontological theories – those which hold that what we ought to do is sensitive to our beliefs about the facts of our situation – is to give a principled and non-problematic explanation for why they agree with this verdict. Smith argues that no such explanation available to them.

Against this, I shall argue that deontological theories do not necessarily face the Problem of Ignorance. In general, deontologists can use a decision-theoretic approach to evaluating the importance of information, even in Smith's target cases in which we have no or mistaken beliefs about our situation.[4]

Part 4.2 presents Smith's argument against deontology. **Part 4.3** argues that deontologists should reject Smith's analysis of their position in favour of a decision-theoretic approach to information gathering. **Part 4.4** compares this approach to the relatively piecemeal approach given by Philip Swenson (2016).[5]

4.2 The Problem of Ignorance

A distinctive and admirable feature of Smith and Swenson's respective discussions of the Problem of Ignorance is that both theorists construct original arguments for their positions, based on novel moral principles that are independently plausible. Notably, however, neither theorist engages with the well-established literature in decision theory on the topic of when an individual ought to gather more information.[6] Instead, they draw conclusions about whether deontology can be adequately extended to cases of incomplete information using piecemeal extensions of existing moral theory. Below, I outline their respective arguments. I argue that their

3. As I discuss below (in Part 4.3.1), this assumption about 'lacking beliefs' plays a crucial role in Smith's argument.

4. Like Smith, I will focus primarily on cases involving no beliefs. The approach I give here also applies to cases involving mistaken beliefs.

5. Philip Swenson, "Subjective Deontology and the Duty to Gather Information," *Ethics* 125, no. October (2016): 257–271.

6. The locus classicus of value of information calculations is found in: I. J. Good, "On the Principle of Total Evidence," *The British Journal for the Philosophy of Science* 17, no. 4 (1967): 319–321. For an accessible introduction, see: Michael D. Resnik, *Choices: An Introduction to Decision Theory* (London: University of Minnesota Press, 1987), pp. 57-59. For novel applications in ethics, see: Jackson, "Decision-Theoretic Consequentialism"; Frank Jackson, "How Decision Theory Illuminates Assignments of Moral Responsibility," chap. 2 in *Intention in Law and Philosophy*, ed. Ngaire Naffine, Rosemary J Owens, and John Williams (Aldershot, UK: Ashgate, 2001), 19–36.

piecemeal nature leads to an unduly narrow – and, in fact, mistaken – approach to determining whether deontologists face the Problem of Ignorance.

4.2.1 Smith's Argument

Smith's contention is that if you are a deontologist then you cannot, by your own lights, have a moral duty to gather more information before acting. If true, then this is a startling conclusion. Given that the critique appears to target all deontological moral theories, it would effectively refute the entire deontological tradition. How does Smith argue for this?

Smith adopts a simple but powerful argumentative strategy: she distinguishes between two jointly exhaustive types of duties that deontologists might posit, and then proceeds to argue that neither type allows deontology to avoid the Problem of Ignorance.

The first type of duty is a *free-standing* one that holds that for *all* cases of uncertainty, you should gather information. Smith rejects any approach of this type:

> Scrutiny of these proposed freestanding duties to seek information reveals that they have several flaws: either they provide poor advice, or they provide insufficiently detailed advice, or they provide advice for which no rationale is forthcoming.[7]

Simply put, any plausible moral theory should be able to distinguish between cases where we should gather information (and to what extent) from other cases where we should not or need not. Since a freestanding approach is (by definition, it seems) insensitive to context, it is difficult to see how it could correctly balance the various considerations of cost, reliability of information, and so on, which are relevant to determining whether we morally ought to gather more information.[8] On this basis, Smith concludes that deontologists cannot avoid the Problem of Ignorance by positing a freestanding duty to gather more information.

If this is correct, then it seems that a deontological theory must instead adopt a *derivative* duty, one that is grounded on the fact that gathering information can help you to discharge other duties you have. Unlike with freestanding duties, it is easier to see how derivative duties will be sensitive to the moral stakes of your situation: if gathering information will help you to better discharge your other duties, then you have duty to do so; if it won't, then you don't.

7. Smith, "The Subjective Moral Duty to Inform Oneself before Acting," p. 19.

8. For a cogent response to this charge, see: Tenenbaum (2017), who holds that we have a defeasible duty to gather information until we know the relevant facts. This requirement is defeasible in that it does not apply in all cases, e.g. those in which gathering more information is not accessible or too costly to access. Tenenbaum, "Action, Deontology, and Risk: Against the Multiplicative Model," p. 20.

However, Smith argues that, due to the particular commitments of deontology, any such duty will also incorrectly determine whether to gather more information. This conclusion is said to follow from the following principles, which I shall call **Subjective Duty** and **No Deontic Value**.[9] I will briefly explain how these premises of Smith's argument help to establish her main result: a case where deontology incorrectly permits us to avoid gathering information.

Premise 1: Subjective Duty

Smith offers the following general principle that deontologists might accept when determining whether or not we ought to gather more information:

Subjective Duty: You have a duty to gather information if and only if you believe that doing so would lead you to produce the maximum amount of deontic value.[10]

I shall later argue that deontologists should reject **Subjective Duty**. For now, however, we will see how it plays an important role in Smith's argument. Applied to **Diagnostics**, **Subjective Duty** appears to entail the following moral code:

Code C

Clause 1. When you believe that you should cure and not-harm the patient, you should:

a. Give Drug A, if you believe doing so will cure the patient, or;

b. Give Drug B, if you do not believe that Drug A will cure the patient.

Clause 2. When you believe that you can gather more information relevant to treating a patient, then you should do so if and only if you believe that doing so would lead you to produce the maximum amount of deontic value.

Smith concludes, however, that **Code C** incorrectly leads you to avoid gathering information once it is coupled with the following assumption: **No Deontic Value**.

9. Smith, "The Subjective Moral Duty to Inform Oneself before Acting," pp. 32-34.

10. This is a simplified version of Smith's principle. For the purposes of the argument given here, nothing of substance is lost by the simplification. Compare: "An agent has a subjective derivative prima facie duty to do what he believes is acquiring information if and only if he believes that doing what he believes is gathering information would leave him subsequently to produce the maximum amount of deontic value (typically through his doing what he then believes to be carrying out the various deontic duties that would then be incumbent upon him)." ibid., p. 24.

Premise 2: No Deontic Value

Smith argues that some actions can *create* a duty that would not otherwise exist. For example, making a promise can create a *pro tanto* duty to perform the promised action, where otherwise this duty would not exist. Likewise, she argues, in cases like **Diagnostics**, your action of investigating or not investigating will create a duty (by Clause 1 of **Code C**) to perform a particular treatment by leading you to believe whether the patient is or is not receptive to Drug A.

Drawing on earlier work, Smith contends that deontology must hold that there is no deontic value to creating a duty, or else we would be required to create increasingly onerous duties: rather than promising to look after someone's goldfish, we would be required to promise to look after their children; we may even be required to damage others' property in order to create the weighty obligation to pay compensation; and so on.[11] Since we are not required to make such promises or, more generally, to create more burdensome duties, Smith concludes that deontologists must accept:

No Deontic Value: There is no positive deontic value to satisfying a created duty, but only a negative value to violating such a duty.[12]

Applied to the present case, Smith argues that since gathering information can create duties (once you become aware of a fact, you can come under a duty to respond to it appropriately), there will be no additional deontic value associated with gathering it as opposed to not doing so.

It is worth noting that deontologists have numerous resources for denying **No Deontic Value**, if they wish. For instance, they might argue that gathering information does not create duties in the relevant sense: rather, gathering information involves *discovering* what your duties actually are.[13] On top of this, deontologists can point out that they can take into account personal costs in determining whether one has an obligation to do what is otherwise best.[14] They can also deny that we should create obligations to compensate others (say, by harming them and becoming duty-bound to compensate), since we have reasons to refrain from harming others, even if our compensating them would bring them greater benefits overall.[15] More generally, there is no reason to think that deontologists hold that the more onerous a moral duty, the more valuable it is. However, it turns out that these lines of argument are unnecessary: as we shall see, **No Deontic Value** is a red-herring that

11. Holly M. Smith, "A Paradox of Promising," *Philosophical Review* 106, no. 2 (1997): 153–196; Smith, "The Subjective Moral Duty to Inform Oneself before Acting"

12. Smith, "The Subjective Moral Duty to Inform Oneself before Acting," p. 29.

13. See: Peter A. Graham, "In Defense of Objectivism about Moral Obligation," *Ethics* 121, no. 1 (2010): p. 91.

14. See: Lazar, "Deontological Decision Theory and Agent-Centered Options."

15. See: Seana Valentine Shiffrin, "Wrongful Life, Procreative Responsibility, and the Significance of Harm," *Legal Theory* 5, no. 2 (1999): 117–148; Thomson, *The Realm of Rights*, pp. 228-48.

Option	*A-Receptive*	*Not A-Receptive*
Drug A	0	-1
Drug B	-1	0

Table 4.1: **No Deontic Value**, applied to **Diagnostics**.

does nothing to block deontology from adopting a decision-theoretic solution to the Problem of Ignorance.

As such, for the sake of argument, we will accept **No Deontic Value** and assign all outcomes in which you fulfil a created duty a value of zero.[16] Without loss of generality, Table 4.1 lists the values for the possible outcomes of your choices in **Diagnostics**.

Conclusion

Smith argues that **Subjective Duty** and **No Deontic Value** jointly entail that you will not be required to gather more information, since gathering information will not lead you to bring about the (uniquely) maximal amount of deontic value, making it morally permissible but not obligatory. As shown in Figure 4.1, whatever you choose, getting more information does not allow you to achieve more deontic value than not getting information. As such, it is permissible for you to choose either option.

More generally, the argument against a derivative duty to gather information can summarised as follows:[17]

Premise 1: You ought to gather more information before acting if and only if you believe that doing so will produce uniquely maximal deontic value. (**Subjective Duty**)

Premise 2: The deontic value of gathering more information is never greater than that of not gathering information. (from **No Deontic Value**)

Conclusion: Therefore, it is never the case that you ought to gather more information before acting.

Thus, having considered both freestanding and derivative approaches to the Problem, Smith concludes that there is no subjective deontological duty that can save deontology from the Problem of Ignorance.

16. It is worth emphasising, however, that talk of 'zero-value' is misleading. As Smith herself notes (note 28, p. 29), given that we are using interval scales, the zero point is arbitrary, since the values can be rescaled up to positive affine transformation without distorting the ordering of the outcomes.

17. See also: Smith, "The Subjective Moral Duty to Inform Oneself before Acting," pp. 31-32.

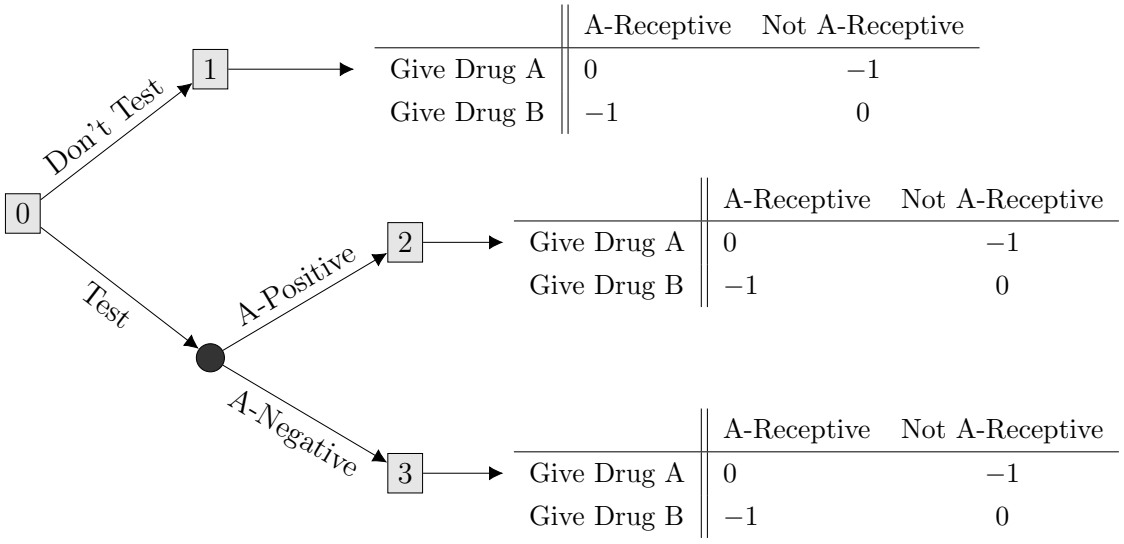


Figure 4.1: Smith's Argument, applied to **Diagnostics**

4.3 A Decision-Theoretic Approach

Subjective Duty determines what you ought to do on the basis of your full beliefs, rather than on your degrees of belief. I will now argue that deontologists should reject **Subjective Duty** because – even in the specific cases that Smith is targeting – we can and should appeal to our degrees of beliefs to determine what we ought to do. I will then show how this approach allows deontologists to avoid the Problem of Ignorance.

4.3.1 Smith's Argument for Subjective Duty

To motivate **Subjective Duty** and perhaps to preclude a decision-theoretic approach to the Problem, Smith makes the following caveat:

> First, I will focus primarily on agents who have false beliefs or no relevant beliefs, rather than agents who are merely uncertain about their prospective action's character. This restriction will not distort the inquiry or its outcome. If a moral theory flunks the test of appropriately handling the duty to gather information for agents having false beliefs or no relevant beliefs, then it is inadequate, even if it passes this test for agents laboring under uncertainty.

This restriction determines the content of **Subjective Duty**, which defines duties in terms of beliefs or lack of belief, rather than in terms of uncertainty. Recall:

Subjective Duty: You have a duty to gather information if and only if you believe that doing so would lead you to subsequently produce the maximum amount of deontic value.

As we saw, this restriction plays a crucial role in Smith's argument. In **Diagnostics**, since (by stipulation) you do not believe that running the test will lead you to produce the maximum amount of deontic value (due to the **No Deontic Value** assumption), it is not the case that you ought to do it.

Unfortunately, Smith never explains what she means by 'belief' or what it means to have 'no relevant beliefs'. More to the point, she never explains why your having no beliefs or false beliefs about a proposition entails that you cannot be more, less or similarly confident that the proposition is true as opposed to its negation. As it stands, there is no apparent inconsistency between not believing that the patient is A-Receptive, not believing that she is not A-Receptive, while nevertheless being more, less, equally, or similarly confident that she is receptive to Drug A as opposed to not receptive to it. This suggests that Smith's target cases – those involving no or mistaken beliefs – and cases involving uncertainty are not mutually exclusive. Indeed, I shall argue that her target cases are a proper subset of cases of uncertainty. If that is correct, deontologists can hold, for familiar decision-theoretic reasons, that your degrees of confidence can make it the case that you should gather information *even if* you do not believe that doing so will maximise deontic value (thus rejecting the 'only if' clause in **Subjective Duty**).

4.3.2 A Deontic Value of Information Approach

In presenting a positive proposal of how deontology can avoid the Problem of Ignorance, I will introduce some terminology that helps to bridge deontological moral theory and decision theory.

For the moment, I will try (where possible) to avoid the term 'duty'. This is because duty-speak can be unclear about how exactly duties are relevant to decision-making in risky situations. For instance, are duties the *output* of a decision process (e.g. your duties are just the set of permissible actions)? Or are duties the *inputs* of a decision process (e.g. you consider your *pro tanto* duties in your deliberations regarding which of your available actions are morally permissible)? Or are duties some *function* from the grounds of obligation to some set of permissible actions (whereby your duty is to consider your reasons for action and to determine, from them, what you morally ought to do)?

In an attempt to side-step this ambiguity, I will adopt the following decision-theoretic framework.[18] A *choice context* is a set of actions, a set of possible states of the world, and a belief function defined over those states that measures how

18. For a general framework for modelling moral theories, one that allows for agent-relative perogatives, menu-dependence, and other structural features commonly associated with deontology, see: Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories."

confident you are that a particular state is actually the way the world is. When an action is performed under a particular state of the world, it produces an *outcome*. This outcome encodes whether moral considerations have been upheld or violated: it may represent the nature of the action (for instance, it may represent the fact that you intentionally lied) and it may also represent the causal consequences of your action (for instance, it may represent whether or not your interlocutor believed your lie). An action's *prospect* is the set of its possible outcomes and their respective probabilities of obtaining. We will assume that these prospects can be ordered and represented by a *deontic value function*, which assigns a higher deontic value to outcomes, the more morally important they are. A *moral decision rule* is a function from a choice context to a set of morally permissible actions. It draws on some or all of the above information to determine what, if anything, is morally permissible.

Understood in this framework, the challenge for deontologists is to find an unproblematic moral decision rule that allows them to avoid the Problem of Ignorance. Smith's claim is that no such moral decision rule exists or – at the very least – coming up with such a rule would involve 'dauntingly complex theoretical issues.'[19]

For example, one reason why Smith denies that deontologists can rely on degrees of belief is that, in cases like the one she is concerned with, modelling an agent's lack of belief in terms of degrees of uncertainty may involve relying on the Principle of Indifference, which is a problematic epistemic rule.[20] Roughly speaking, the Principle of Indifference says that when we lack any evidence about our situation, we should assign equal probabilities to the various possibilities that might occur. If you do not know whether the patient is A-Receptive or not, then you should have a 0.5 credence in either possibility. It is well known, however, that the Principle of Indifference faces the problem that possibilities can be redescribed in seemingly equally eligible ways to yield different numbers of possibilities, leading to a conflicting assignment of probabilities.[21]

In response, however, note that even if the Principle of Indifference is false (which is debatable), deontologists do not need to invoke it; they can use other epistemological tools.[22] Indeed, orthodox Bayesian epistemology holds that, if we are rational, we always have precise and coherent degrees of belief even if we do not use the Principle of Indifference. The mere fact that we lack all-out belief – and that the Principle

19. Smith, "The Subjective Moral Duty to Inform Oneself before Acting," p. 33.

20. This principle was so-named by John Maynard Keynes, *A Treatise of Probability* (London; New York: Macmillan; AMS), ch. 4. For Smith's discussion, see: Smith, "The Subjective Moral Duty to Inform Oneself before Acting," p. 33.

21. For a classic discussion of this issue, see: Bas van Fraassen, *Laws and Symmetries* (Oxford: Clarendon Press, 1989), ch. 12.

22. For a recent defence of the Principle of Indifference and summary of other approaches, see: Richard Pettigrew, "Accuracy, Risk, and the Principle of Indifference," *Philosophy and Phenomenological Research* 92, no. 1 (2016): 35–59.

of Indifference is controversial – does not entail that we do not have well-defined degrees of belief.

Perhaps, however, Smith's target cases are not simply those in which we lack all-out belief; they are cases where we lack all-out belief because we lack evidence. It is arguable – though still a matter of significant debate – that lacking evidence prevents us from having precise degrees of belief.[23]

However, even if such cases present problems for orthodox Bayesianism, it turns out that there are well-established, alternative Bayesian approaches to dealing with them. Specifically, deontologists can represent themselves as having an indeterminate credal state that reflects their equivocal, incomplete, or otherwise ambiguous evidence.[24] Where a determinate degree of belief specifies a specific numerical value (indeed, precise up to infinite decimal places), an indeterminate degree of belief admits of a range of such values, modelled by a set of probability functions. By positing sets of probability functions, deontologists can model the cases that Smith targets, without appealing to the Principle of Indifference.[25]

Adopting this approach in **Diagnostics**, we will model your epistemic situation of 'lacking beliefs' in terms of indeterminate degrees of belief. In this case, let's suppose that you are roughly equally confident that the patient is A-Receptive rather than not so, such that the probabilities range within the interval [0.4, 0.6] (or ≈ 0.5, for ease of exposition).[26]

With this prior, albeit indeterminate, degree of belief, running the test will allow you to update your degrees of belief and choose from a more informed standpoint. Let us suppose that the test is known to be 99% accurate. Given this information, being a rational agent you will update your degrees of belief to reflect the evidence. As we shall see in **Diagnostics**, this information will improve the prospects associated with giving one drug rather than the other. Following this approach, de-

23. For a defence of indeterminate credences, more generally, see: James M. Joyce, "A Defense of Imprecise Credences in Inference and Decision Making," *Philosophical Perspectives* 24 (2010): 281–323; Alan Hájek and Michael Smithson, "Rationality and Indeterminate Probabilities," *Synthese* 187 (2012): 33–48; Isaac Levi, "Why Indeterminate Probability is Rational," *Journal of Applied Logic* 7, no. 4 (December 2009): 364–376. For an argument in favour of the orthodox view, see: Elga, "Subjective Probabilities should be Sharp."

24. James M. Joyce, "How Probabilities Reflect Evidence," *Philosophical Perspectives* 19 (2005).

25. For defences of this approach, see: Isaac Levi, *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance* (Cambridge, Massachusetts; London, England: MIT Press, 1980); Jeffrey, *The Logic of Decision*; Mark Kaplan, *Decision Theory as Philosophy* (Cambridge: Cambridge University Press, 1996); Joyce, "How Probabilities Reflect Evidence"; Joyce, "A Defense of Imprecise Credences in Inference and Decision Making."

26. Note, however, that we can make the interval much wider: indeed, potentially maximally so – representing, perhaps, your complete and utter incomprehension of your situation. New evidence will have much less of an effect on your credences in such cases, but may still improve the prospects of your treatment. As such, it may be morally required. For discussion, see: Joyce, "A Defense of Imprecise Credences in Inference and Decision Making"; Aron Vallinder, "Imprecise Bayesianism and Global Belief Inertia," *British Journal for the Philosophy of Science* 0, no. June (2018): 1–26.

ontologists should reject **Subjective Duty** in favour of the following moral decision rule:

Choose Undefeated Prospects: An action is morally permissible if and only if its prospect is undefeated by that of any alternative action.

An 'undefeated prospect' is one that is not ranked lower than any other prospect. For current purposes, we do not need to fully define the properties that govern whether one prospect is ranked higher than another; we can simply hold that, all else equal, a treatment with a higher probability of better outcomes defeats a treatment with a lower probability of better outcomes.[27] This can be conveniently represented in terms of one action having a strictly higher probability-weighted average (or 'expected') deontic value (EDV) than another.[28]

Choose Undefeated Prospects solves the Problem of Ignorance because it allows deontologists to use what is equivalent to a 'value of information' calculation to determine when it is morally required to gather more information. Roughly put, value of information calculations involve comparing your optimal future actions in an informed context (which results from having gathered more information) versus an uninformed one. If the informed context has an optimal action whose prospect defeats that of every action in the uninformed context, then you should choose to gather more information.[29]

For example, in **Diagnostics**, you have a choice between acting in an informed choice context (by running the test) or in an uninformed choice context (by not running the test). The rule **Choose Undefeated Prospects** determines whether or not you should run the test by calculating whether the optimal action in the informed choice context has higher expected deontic value than the optimal action in the uninformed choice context. Figure 4.2 illustrates your choice situation based on your prior and posterior degrees of belief, giving the expected deontic values of your available courses of action.

27. As mentioned in Chapter 1, a defeat relation is weaker than a weighing relation. Unlike a weighing relation, a defeat relation can allow for incommensurable considerations, strong moral dilemmas, (arguably) moral options, and even intransitivity. See: Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories," p. 432.

28. Note that this expected value representation can be weakened in several respects, notably by dropping the Completeness and Continuity axioms, allowing for incommensurable considerations and a strong version of lexical priority, respectively: see Robert J. Aumann, "Utility Theory without the Completeness Axiom," *Econometrica* 30, no. 3 (1962): 445–462; Hausner, "Multidimensional Utilities." Indeed, my argument can be run without expected value altogether, instead using a barer choice-theoretic approach based on a defeat relation, as described in: Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories."

29. For an accessible introduction, see: Resnik (1987). For details on how such calculations work with indeterminate probabilities (as in our current case), see: Bradley and Steele (2016). Resnik, *Choices: An Introduction to Decision Theory*; Seamus Bradley and Katie Steele, "Can Free Evidence Be Bad? Value of Information for the Imprecise Probabilist," *Philosophy of Science* 83, no. 1 (2016): 1–28.

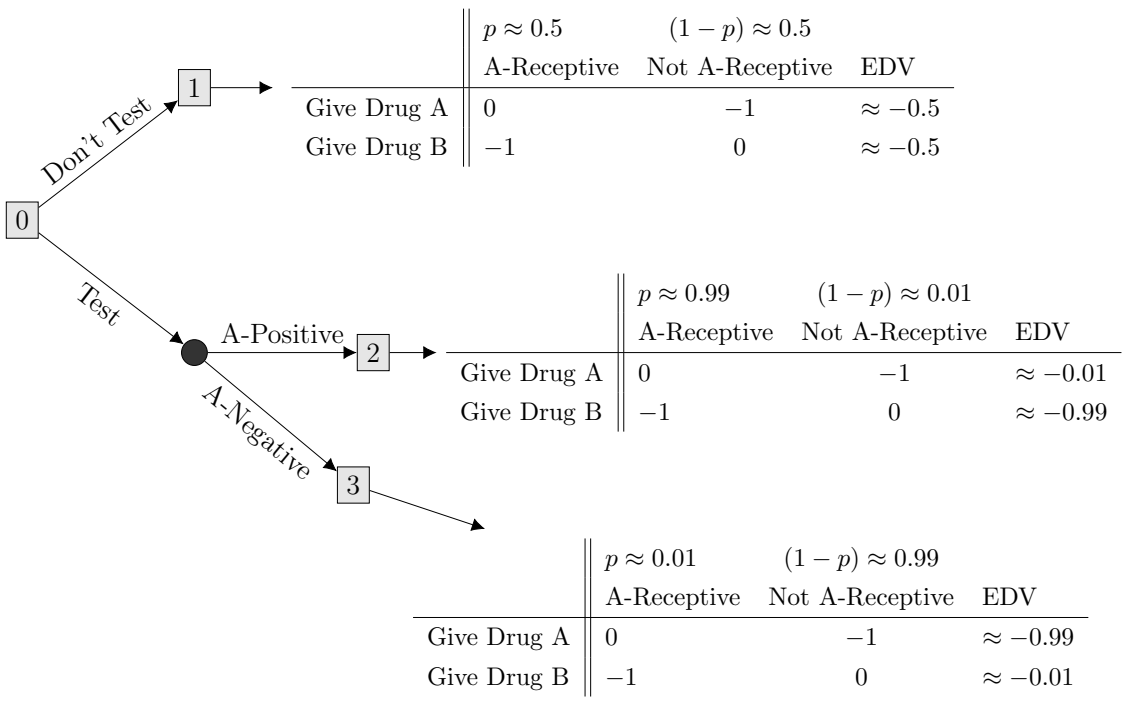


Figure 4.2: Value of Information Calculation

If you choose to remain uninformed, the morally optimal action of your later choice situation has an expected deontic value of between -0.4 and -0.6 (or ≈ -0.5 for short, as in the diagram above). Compare this with the context in which you run the diagnostic test. If you Test, then the optimal actions have an expected deontic value of approximately -0.01, which is greater than that of the optimal actions in the uninformed context.[30] According to **Choose Undefeated Prospects**, you are morally required to run the diagnostic test. This result shows that there is a subjective duty (or moral decision rule) that deontologists can adopt to avoid the Problem of Ignorance.[31] The question, then, is whether deontologists can avail themselves of this solution without giving up their core substantive commitments.

4.3.3 Objections

Perhaps anticipating this approach, Smith offers some lines of resistance to any decision-theoretic model of deontology. For instance, she mentions that deontologists are not obviously committed to the idea that deontic value is representable by a linear value function.[32] Some deontologists agree.[33] As Sergio Tenenbaum puts it:

> There is no obvious measurable value that the deontologist is trying to bring

30. Strictly speaking, this will be an interval of values around the value 0.1.
31. Note also that this approach can be extended to account for cases of multiple conflicting duties, costly or not fully reliable information, or uncertainty about our future actions.
32. Smith, "The Subjective Moral Duty to Inform Oneself before Acting," p. 33.
33. See, for example: Aboodi, Borer, and Enoch, "Deontology, Individualism," p. 272.

> about to the greatest degree. So when a deontologist multiplies a probability by a value, it is far from clear what this product represents. If I am bound by a rule that prohibits lying because lying expresses disrespect, does raising the probability of lying always express disrespect (rather than at most incurring a risk of expressing disrespect)? And if risking disrespect is a form of disrespect, is it disrespectful in proportion to the risk?[34]

There are a few ways of understanding this concern. On one interpretation, the concern is that deontological moral considerations are not ordered in the right way to allow for a well-defined measure of deontic value. Bracketing the fact that Smith's argument presumes that deontological considerations *can* be represented by a deontic value function, there are reasons to think that deontology violates at least some of the axioms of orthodox expected value theory.[35] However, this requires further investigation, not least because decision theorists have made many extensions and generalisations of orthodox decision theory that allow it to accommodate risk attitudes, incommensurability, and lexical priority – all while maintaining a systematic and coherent approach to risk.[36] It is therefore not enough to show that deontology cannot be represented by *some* particular version of expected value theory; those pressing this approach need to show that deontology cannot be represented by *any* version of expected value theory. We are yet to see an impossibility result to this effect.

A weaker interpretation of the concern attends to the question of whether deontologists should be *risk neutral*: that is, whether they should consider two prospects that have same expected deontic value to be equally choice-worthy, even if one prospect has much greater variance (or greater 'spread') in possible deontic value than the other. As it turns out, however, non-neutral attitudes to risk are consistent with orthodox decision theory.[37] Although deontologists are yet to settle such questions as whether non-neutral risk attitudes are morally appropriate, the solution given above is compatible with various answers to the question.[38]

34. Tenenbaum, "Action, Deontology, and Risk: Against the Multiplicative Model," p. 9

35. For arguments to this effect, see: Frances M. Kamm, "Supererogation and Obligation," *Journal of Philosophy* 82, no. 3 (1985): 118–138; Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, ch. 7; Lazar, "Deontological Decision Theory and Agent-Centered Options"; Lazar, "Limited Aggregation and Risk"; Tenenbaum, "Action, Deontology, and Risk: Against the Multiplicative Model."

36. Pratt, "Risk Aversion in the Small and in the Large"; Aumann, "Utility Theory without the Completeness Axiom"; Hausner, "Multidimensional Utilities." See also: Edward Elliott, "A Representation Theorem for Frequently Irrational Agents," *Journal of Philosophical Logic* 46, no. 5 (2017): 467–506; Dietrich and List, "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories."

37. Pratt, "Risk Aversion in the Small and in the Large"; Arrow, "The Theory of Risk Aversion"; Broome, *Weighing Goods*, Ch. 4.

38. Given that **Choose Undefeated Prospects** is logically weaker than expected value theory, one alternative approach for deontologists might be to adopt Lara Buchak's Risk-Weighted Expected Utility Theory. See: Buchak, *Risk and Rationality*. One worry with this particular approach, however, it that is might lead deontologists back to a version of the Problem of Igno-

For their part, deontologists might worry that this proposal is too consequentialist: it determines the deontic value of gathering information based on the deontic value your later choices, which in this case amount to correctly or not correctly treating the patient, based on your evidence. To anyone with deontological inclinations, it seems that your choosing to gather more information is based on *more* than just achieving the best consequences for your patient – there is something *in itself* morally important about your efforts. For instance, gathering more information can be a way of showing concern and respect for your patient. The objection goes: if the solution I've presented here does not accommodate that kind of non-consequentialist consideration, then it is not an appropriate model of deontology.

As others have noted, this type of consideration can be included in a decision-theoretic model like the one above.[39] In Smith's terms, the duty to gather information would be based on both a free-standing *and* derivative duty. The free-standing duty can be represented by giving some additional positive deontic value α to any course of action that includes gathering information out of concern and respect for others.[40] The derivative duty is represented by the difference in expected deontic value that you gain from gathering information. This approach is entirely consistent with the decision-theoretic model presented earlier, and does seem to offer a more accurate representation of deontological concerns.

Overall, while much work remains to be done in constructing a fully-fledged deontological decision theory, it appears that deontologists can adopt **Choose Undefeated Prospects** to avoid the Problem of Ignorance.

4.4 Comparison with Swenson's Account

Let us now compare the approach given above with that given by Philip Swenson (2016).[41] In response to Smith's argument, Swenson accepts both **Subjective Duty** and **No Deontic Value**. However, he denies that **No Deontic Value** applies to actions that create duties by gathering more information. In **Diagnostics**, this would allow him to assign higher deontic value to gathering more information as opposed to not doing so.

What is the rationale for treating information gathering in this way? Swenson argues that deontologists should be concerned with best *approximating* deontic ideals, and that gathering information is a useful way of better achieving this. Thus, in

rance, whereby individuals are not only led to reject free, reliable information and to also choose a state-wise dominated course of action. See: R.A. Briggs, "Costs of abandoning the Sure-Thing Principle," *Canadian Journal of Philosophy* 45, nos. 5-6 (2015): 827–840.

39. Colyvan, Cox, and Steele, "Modelling the Moral Dimension of Decisions."

40. This could be a constant value or some function of other morally relevant considerations. For present purposes, we do not need to go into detail about this.

41. Swenson, "Subjective Deontology and the Duty to Gather Information."

Diagnostics, you should presumably run the test because this best approximates the more important deontic ideal of treating the patient correctly.

More generally, he posits a number of additional 'approximation principles' that, he argues, deontologists would find appealing. For instance, he posits a *Subjective Approximation Principle for Promise Keeping*, a *Subjective Approximation Principle for Justice in Layoffs*, and so forth. Each of these rules roughly holds that we should perform an action that best approximates (in some undefined sense) the particular deontic ideals at play.

Thus, on Swenson's account, deontic value is not only a measure of the significance of a deontic ideal; it is also a measure of how well an action approximates some set of deontic ideals. This is given by his Deontic Value Principles:

First Deontic Value Principle: All things being equal, Option A has higher deontic value than Option B if A approximates some deontic ideal to a greater degree than B.

Second Deontic Value Principle: All things being equal, Option A has higher deontic value than Option B if A achieves some deontic ideal X and X is (in the relevant context) a more significant ideal than any deontic ideal achieved by B (where achieving a deontic ideal counts as maximally approximating it).

More generally, Swenson's strategy proposes that deontologists should respond to the Problem of Ignorance by positing additional principles on a case-by-case basis. As he puts it:

> When Subjective Deontologists are confronted with cases in which it is intuitive that the agent has a duty to gather information, they should look for principles that allow them to assign deontic value in a manner that accounts for this duty. Given that the agent has the duty, the failure to gather information will be regrettable. This reveals that the deontic ideal has not been achieved. So the very fact that the agent has the duty is evidence that the deontic values should be assigned in a way that accounts for the duty. Thus it is highly plausible that there will always be (perhaps very complex) principles of deontic value that can do the necessary work.

Swenson is optimistic that deontology will be able to construct a general and consistent set of principles to cover all cases involving incomplete information. However, I am less confident about this approach. This is because, as it stands, Swenson's proposal is underspecified in a number of important respects. For instance, we do not yet know why the regret heuristic will induce a well-defined measure of deontic value.[42] Nor do we know what it means for an action to *approximate* an ideal, and why this notion would support a well-defined measure of deontic value.

42. For instance, in cases where there is no fact of the matter what would have occurred, had we done otherwise, it seems that the regret heuristic might lead us to assign higher deontic value to

To illustrate this concern, consider the deontic ideal of correctly treating your patient.[43] It seems that there are multiple ways of approximating this ideal. Suppose that you have the following options:

Send to Dr Knowledgeable: Refer your patient to a doctor who is less caring than you (e.g. worse bedside manner), but more knowledgeable than you about the most appropriate treatment plans available.

Send to Dr Caring: Refer patient to a doctor who is more caring (e.g. better bedside manner), but less knowledgeable than you about the most appropriate treatment plans available.

Send to yourself: Treat the patient yourself, giving a normal amount of care and competence.

Whatever you choose, there will be some other option available that better approximates the deontic ideal in some respect. Suppose that you begin by not referring. Then you (rightly) become concerned that you are potentially depriving your patient of more effective treatments, so you refer the patient to Dr Knowledgeable. However, Dr Knowledgeable's horrible bedside manner makes you quickly redirect the patient to Dr Caring. However, as good as Dr Caring's intentions are, his treatment plan is probably not going to be the most appropriate, so you finally send the patient back to yourself. Thus, you've led yourself (and your patient) in a circle. Each option better approximates *in some respect* the deontic ideal of correctly treating your patient. However, by that same token, each option will also be regrettable in some respect. Without further information about how the approximation relation is to be understood, it is possible that the relation may sometimes be intransitive, preventing us from assigning well-defined deontic values to options.

This is merely to illustrate that Swenson's account does not explain how its most critical aspects function as he hopes, and so it is unclear why we should expect our adding more case-by-case principles will lead to an informative and internally-consistent account. Without a clearer sense that the foundations of the account are secure, deontologists should be wary of building further complex principles on top of them.

---

whatever we choose *not* to do. See: Caspar Hare, "Obligation and Regret When There is No Fact of the Matter About What Would Have Happened if You Had not Done What You Did," *Noûs* 45, no. 1 (2011): 190–206.

43. The following is a similar case to those found in: Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. The best explanation for such cases seems to me to be given by Luce and Raiffa (1957), who observe that intransitive choices can arise "when a subject forces choice between inherently incomparable alternatives". See: R. Duncan Luce and Howard Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York: John Wiley / Sons, Ltd., 1957). See also: John Cusbert, "Acting on Essentially Comparative Goodness," *Thought: A Journal of Philosophy* 6, no. 2 (2017): 73–83.

Another problem is that it is entirely unclear how Swenson's account accommodates degrees of uncertainty. Recall that in **Diagnostics**, you are 99% certain that the test is accurate. Should you run the test? Swenson's account has no answer. Indeed, it falls silent for *any* degree of uncertainty, since its approximation and deontic value principles are formulated without any reference to degrees of uncertainty. The question is: can it be extended to cover such cases, thus providing a general solution to the Problem of Ignorance for deontologists?

Following Swenson's preferred approach of positing additional (potentially complex) principles, we might attempt to extend Swenson's account to cases of uncertainty by holding that there is a close connection between uncertainty and approximation. For example, the following would be a natural extension of his account:

Third Deontic Value Principle: All else equal, Option A better approximates a deontic ideal X than Option B when A is more certain to achieve the ideal than B.

Unfortunately, this cannot be the correct approach. 'Achieving' a deontic ideal, on Swenson's proposal, "counts as maximally approximating it".[44] So, according to the Third Deontic Value Principle, Option A better approximates a deontic ideal by virtue of being more certain to better (indeed, maximally) approximate it. This is clearly circular. Hence, there must be a different sense of approximation at play: perhaps, a subjective one. Accounting for this requires an additional principle, such as:

Subjectively Approximate Maximum Approximation: All else equal, if Option A subjectively approximates maximum approximation of a deontic ideal more than Option B, then you should choose Option A rather than Option B.

By virtue of being more certain that running the test will lead you to better approximate the deontic ideal of treating your patient, you ought to do so. Note, however, that even this is not enough. What if all else is *not* equal? What are we to say about cases where the two dimensions of deontic value (approximation and importance) disagree? For instance, what if one option has a low probability of achieving maximum approximation, compared to another option where there is a high probability of achieving lesser approximation? Until we know more about the core aspects of Swenson's account, we cannot treat it as a general solution to the Problem of Ignorance.

Of course, an easy alternative approach would keep uncertainty separate from approximation. To deal with cases of uncertainty, we might posit a moral decision

44. Swenson, "Subjective Deontology and the Duty to Gather Information," p. 264.

rule that takes into account degrees of uncertainty and deontic value. However, it is unclear how this will not just turn out to be a heavily cloaked version of the decision-theoretic solution presented earlier.

4.5 Conclusion

The Problem of Ignorance does not apply to deontological moral theories that are compatible with expected value theory. Using value of information calculations, deontologists can determine when we morally ought to gather more information before acting. However, to determine whether deontologists can truly avail themselves of this solution, we must further investigate the compatibility of deontological commitments with the core axioms of expected value theory. Nevertheless, as it stands, there are reasons to be optimistic that deontology can guide us through an uncertain world.

Chapter 5

Authority, Obedience, and Uncertainty

> There is but one power to which I can yield a heart-felt obedience, the decision of my own understanding, the dictate of my own conscience.
>
> William Godwin,
> *An Enquiry Concerning Political Justice* (1793)

5.1 Introduction

Suppose that you are serving in the military, and have just been ordered by your superior officer to commandeer a civilian's van to deliver live-saving goods to a distant village.[1] You have doubts about whether this is the best course of action. Moreover, as it happens, you are certain that if you disobey, it will go undetected. What ought you to do?

Many hold that the answer depends at least partly on whether your superior officer is a *legitimate* authority. What does it take to be a legitimate authority? According to Joseph Raz's Service Conception, an authority is legitimate just in case it reliably helps us to better act in accordance with the reasons that apply to us. This idea is more precisely captured by:

The Normal Justification Thesis: An authority is legitimate if and only if obeying its orders would lead you to better conform with the reasons that independently apply to you, compared to if you acted on your own assessment of

1. This case is loosely based on: Raz, *Practical Reason and Norms*, pp. 38-45; Joseph Raz, *From Normativity to Responsibility* (Oxford; New York: Oxford University Press, 2011), p. 116.

those reasons.[2]

Let us assume that in the case at hand, you have most reason to save as many lives as you can, even if doing so involves violating another's property rights. However, if you can save just as many lives without violating anyone's property rights, then you should take the less rights-infringing course of action. According to the Normal Justification Thesis, if obeying your superior's orders is a more reliable means of helping you to save more lives compared to following your own judgement, then your superior officer has legitimate authority over you.

To better understand the distinctive normative power of legitimate authority, we need to know more about the kind of obedience such authorities can demand of us. This is defined by:

The Pre-Emption Thesis: when a legitimate authority issues you a command, it thereby gives you a:

Content-Independent Reason: A reason to do as it commands *because a legitimate authority commanded it*, not because of the content of the command.[3] And an:

Exclusionary Reason: A reason to *not act upon* your own assessment of your situation.[4]

Applied to the present case, the Pre-Emption Thesis holds that if your superior officer is a legitimate authority, you should obey by not acting on your own ideas of how best to deliver the goods to the village, and also by doing as commanded simply because it was directed by a legitimate authority. Obedience in this *pre-emptive sense* is taken to be the correct response to a legitimate authority. Indeed, according to Raz:

2. This *if and only if* formulation attempts to take into account both the Independence Thesis and the Dependence Thesis. The Independence Thesis states that: "the matters regarding which the first condition is met are such that with respect to them it is better to conform to reason than to decide for oneself, unaided by authority" Joseph Raz, "The Problem of Authority: Revisiting the Service Conception," *Minnesota Law Review* 90, no. 1979 (2006): p. 1014. Arguably, however, this is already captured in the Normal Justification Thesis, since in the relevant cases capture by the Independence Thesis are those in which you have most reason to *not* defer to others' judgements and to instead decide on the basis of your own. The Dependence Thesis states that: "All authoritative directives should be based, among other factors, on reasons which apply to subjects of those directives and which bear on the circumstances covered by the directives." (Joseph Raz, *Ethics in the Public Domain* (Oxford: Clarendon Press, 1994), p. 214) This is invoked in the above formulation of the Normal Justification Thesis through the clause that the relevant reasons are those that *independently* apply to you.

3. On this notion, see: Stefan Sciaraffa, "On Content-Independent Reasons: It's Not in the Name," *Law and Philosophy* 28, no. 3 (2008): 233–260.

4. Raz, *The Morality of Freedom*.

> The pre-emptive force of [legitimate] authority is part and parcel of its nature. It cannot succeed as an authority (i.e., succeed in improving our conformity with reason) if it does not preempt the background reasons.[5]

With this theory in hand, let us return to our central case: should you take the van or not? It seems that even though you have doubts about the correctness of the order, if you are sure that your superior officer has more reliable judgement than you, then you should obey.

However, what if you are *uncertain* about whether your superior officer is more reliable than you? Should you pre-emptively obey, despite your doubts? If so, how certain do you need to be that the authority is not legitimate before you should no longer do so? As we shall see, existing answers to these questions are problematic.

Before proceeding, however, it is worth pausing to remember why we should care about how theories of legitimate authority operate in cases of uncertainty. Anyone with a passing knowledge of human history will notice that the practical stakes are high: many of humanity's greatest moral atrocities were the result of obedience to illegitimate authority. Without a clear moral standard for justifiable obedience to authorities of questionable legitimacy, individuals who execute the wrongful commands of authority can almost always assert that they did not *know* that the authority is illegitimate, and that they were "just following orders".[6]

There are also important theoretical stakes at play. In particular, many believe that the Service Conception refutes Philosophical Anarchism, a view which holds that we should never surrender our will or judgement to that of another, as doing so would violate our moral obligation to be autonomous and rational decision-makers.[7] According to the Anarchist, theories of legitimate authority are mistaken because they permit – and, in some cases, require – us to submit our will in this way.[8] To authority, and particularly to political authority, the Anarchist says:

> It is yours to shackle the body and restrain our external actions; that is a restraint we understand. Account your penalties; and we will make our election of submission or suffering. But do not seek to enslave our minds. Exhibit your force in its plainest form, for that is your province; but seek not to inveigle and mislead us. Obedience and external submission is all you are entitled to claim; you can have no right to extort our deference, and command us not to see, and disapprove of, your errors.[9]

5. Raz, "The Problem of Authority: Revisiting the Service Conception," p. 1019.

6. On the jurisprudential difficulties of such cases, see: Mark J. Osiel, "Obeying Orders: Atrocity, Military Discipline, and the Law of War," *California Law Review* 86, no. 5 (1998): 943.

7. See, for example: Leslie Green, *The Authority of the State* (Oxford: Clarendon Press, 1988); Shapiro, *Authority*.

8. William Godwin, *Enquiry Concerning Political Justice and Its Influence on Morals and Happiness* (Toronto: University of Toronto Press, 1946); Robert Paul Wolff, *In Defense of Anarchism* (University of California Press, 1970).

9. Godwin, *Enquiry Concerning Political Justice and Its Influence on Morals and Happiness*, pp. 236-237.

In response, the Service Conception is taken to show that there is no necessary conflict between autonomy, rationality, and obedience to legitimate authority. Firstly, pre-emptive obedience does not require us to not assess the merits of an order or legitimacy of an authority; it merely requires that we not act upon our less reliable judgement of the situation.[10] Secondly, since authorities are legitimate only insofar as they promote our conformity with our reasons for action, far from being in conflict with rationality and autonomy, pre-emptive obedience to legitimate authority is the *very definition* of how rational and autonomous decision-makers respond to the commands of a legitimate authority.[11] In Leslie Green's words, the Service Conception explains how we can act *contrary to the balance of reasons* without thereby acting *contrary to reason.*[12]

As I shall argue, however, cases involving uncertainty threaten to show that the Service Conception requires irrational obedience to authority, thereby vindicating the Anarchist. I propose that decision-theoretic modelling helps the Service Conception to deal with such cases, thereby supporting it as a response to the Anarchist and as a guide to evaluating the justifiability of obedience to authorities of questionable legitimacy.

Part 5.2 discusses existing attempts to deal with uncertainty regarding an authority's legitimacy. I argue that the Service Conception can – and, indeed, should – be modelled in terms of maximising expected value. **Part 5.3** presents a decision-theoretic model of the Service Conception, giving a precise account of when pre-emptive obedience is justified, and when it is not. **Part 5.4** responds to Scott Shapiro's (2004) objections to decision-theoretic approaches to obedience to authority. **Conclusion** follows.

5.2 Uncertain Legitimacy

How should we respond to authorities when we are unsure whether they are legitimate? A natural answer is to hold that, in cases of uncertainty, we should exclude our judgement to the extent that we believe the authority is legitimate. Unfortunately, this proposal conflicts with the very motivations for, and idea of, exclusionary reasons. The motivation for exclusionary reasons is that they save us from costly or unreliable deliberations. The idea of exclusionary reasons is that they have absolute priority over the reasons they exclude.[13] If, upon receiving an order, you were to

10. Joseph Raz, "Legitimate Authority," in *Philosophical Law*, ed. Richard Bronaugh (Westport, Conn.: Greenwood Press, 1978), 3–27; Joseph Raz, "Authority and Justification," *Philosophy and Public Affairs* 14, no. 1 (1985): 3–29.

11. Raz, "Authority and Justification," p. 29.

12. Green, *The Authority of the State*, p. 37

13. As Raz puts it: "Exclusionary reasons always prevail, when in conflict with [the reasons they exclude]." Raz, *Practical Reason and Norms*, p. 40.

partially exclude your judgement by weighing the excluded reasons against the exclusionary reason, then you would be undertaking the costly task of combining your potentially less reliable judgement with that of the potentially legitimate authority. In cases of uncertainty, the 'service' given by legitimate authorities would therefore be undermined.

Perhaps for this reason, Raz rejects the idea of 'partial exclusion'.[14] Instead, he appeals to the idea of *knowability*, such that if we cannot know that an authority satisfies the Normal Justification Thesis by a reasonable inquiry (roughly, one that is worth the cost, given the stakes of the situation), then the authority is not legitimate. This is because, according to Raz:

> [G]enerally speaking, the only reliable way of conforming to authority is through having a reliable belief that it is an authority, and therefore should be obeyed ... When reasonable inquiry will not reveal the case for authority, that case, if it exists at all, is unknowable. It follows that people are not subject to any authority regarding those matters.[15]

This sets out a necessary condition for determining when pre-emptive obedience is justifiable. However, it leaves a lot unsaid. For example, how reliable do our beliefs need to be? What if we are unsure whether we have reliable beliefs about the legitimacy of an authority? Or, what if we are uncertain about whether a reasonable inquiry will give us reliable beliefs about the legitimacy of an authority? Perhaps in an attempt to side-step such matters, Raz qualifies the knowability constraint by assuming that that if we can form reliable beliefs about an authority's legitimacy, then we can also know that these beliefs are reliable.[16] This amounts to applying an additional knowability constraint on the knowability constraint: an authority is legitimate only if we can know (by a reasonable inquiry) that we can have knowably reliable beliefs about its legitimacy.

One problem with this approach is that unless knowledge is perfectly *luminous* (that is, when you know, you *know* that you know), it seems that knowability will not be sufficient for determining whether pre-emptive obedience is justified. And, as it turns out, it seems that knowledge is *not* perfectly luminous.[17] As such, it is unlikely that applying layers of knowability will ever amount to a complete account

14. See Raz's discussion of whether exclusionary reasons are, as Stephen Perry argues, 'reweighting reasons'. See: Stephen Perry, "Second-Order Reasons, Uncertainty and Legal Theory," *California Law Review* 62 (1988): 913–994; Joseph Raz, "Facing Up: A Reply," *California Law Review* 62 (1988).

15. Raz, "The Problem of Authority: Revisiting the Service Conception," p. 1025. See also: Raz, *From Normativity to Responsibility*.

16. Raz, "The Problem of Authority: Revisiting the Service Conception," p. 1025.

17. See: Timothy Williamson, *Knowledge and Its Limits* (Oxford Scholarship Online, 2003), ch. 4 On this point with respect to Raz's more general theory of reasons, see: Ruth Chang, "Raz on Reasons, Reason, and Rationality: On Raz's From Normativity to Responsibility," *Jerusalem Review of Legal Studies* 8, no. 1 (2013): p. 208.

of rational obedience under uncertainty. From a measurement theory perspective, this should be expected: knowability seems to be a binary notion – you can know something or you cannot – whereas uncertainty admits of degrees; as such, it seems that any knowability constraint will need to draw a bright line between shades of uncertain facts to distinguish those which are knowable from those which are not. In doing so, any such constraint seems liable to make incorrect or arbitrary judgements about whether particular uncertain facts should guide us. To avoid these problems, we need more than knowability.

For this reason, it is tempting to build uncertainty *into* the standard of legitimacy, in the following way:

The Normal Justification Thesis (More Likely): You should treat an authority's commands pre-emptively if and only if doing so would *more likely* lead you to better conform with the reasons that independently apply to you, compared to if you acted on your own assessment of those reasons.[18]

If successful, then this approach provides an elegant answer to the problem of what to do in cases of uncertainty (here we will assume that the relevant interpretation of 'likelihood' is informed by your evidence).[19] As we shall see, however, taking this approach leads us to make a substantive commitment about reasons for action. Specifically, it requires us to hold that the normative weight of reasons is *probability-sensitive*. To see why, suppose that reasons are *not* probability-sensitive, and consider the following case: [20]

Questionable Orders: You have been ordered to deliver the life-saving goods by hang glider. If you obey this order, then there is a probability of 0.01 that you will save 101 people, and a probability of 0.99 that you will save no-one (say, because the goods will almost certainly be destroyed when you drop them from such a great height). If you follow your own judgement, you will take the nearby truck, which is slightly slower, but more reliable. Using it has a probability of 0.99 that you will save 100 people, and a probability of 0.01 that you will save no-one (say, by breaking down). All other considerations are equal.

In this case, on a probability-insensitive interpretation of reasons, the Normal Justification Thesis is satisfied because you are *more* likely (probability of 0.01 vs probability of 0) to *better* conform to reason (101 lives saved vs. 100 lives saved) by

18. For instances of this formulation, see: Raz, "Authority and Justification"; Raz, "Facing Up: A Reply."

19. c.f. Raz, *From Normativity to Responsibility*, p. 115.

20. Jackson, "Decision-Theoretic Consequentialism"; Raz, *From Normativity to Responsibility*, pp. 120-128.

pre-emptively obeying. This is problematic because, clearly, obedience is morally reckless! This suggests that the Normal Justification Thesis (More Likely) should not be interpreted in terms of probability-insensitive reasons.[21]

Now, of course, it is possible that you could get lucky and save the 101 individuals. However, the Anarchist would (rightly) chalk this up as a victory, since this kind of indifference to the consequences of obedience makes us "the ready tool of injustice, cruelty, and profligacy; and, if at any time [we] are not employed in their purposes, it is the result of accident, not of [our] own precaution and honesty."[22] In general, any adequate theory of legitimate authority must not tell us to 'just follow orders' and recklessly turn ourselves over to authority.

In technical terms, the probability-insensitive interpretation of the Normal Justification Thesis gives the wrong result because it grants legitimacy and dictates obedience merely on the basis of an ordinal test: it merely assesses whether the authority improves your conformity with reason. Under conditions of certainty, this test is sufficient to guide action. However, to guide action under conditions of risk and uncertainty, we need to also take into account cardinal information: specifically, *how much* improved conformity with reason should we expect from following the authority's orders? Does this potential improvement outweigh the risk that, if the authority is mistaken, we may do much worse if we follow the order? Thus, if we build uncertainty into the standard of justification, we must ensure that we include the relevant cardinal information. This can be done by making the weight of reasons sensitive to probability.[23]

Allowing that the weight of reasons is probability-sensitive yields the following version of the Normal Justification Thesis:

The Normal Justification Thesis (Probability-Weighted): You should treat an authority's commands pre-emptively if and only if doing so would *more*

21. Some might object that this is a case where the authority is clearly mistaken and, as such, the Service Conception allows you to disregard the order (Raz, 1986, p. 62). In response, however, this is not the relevant kind of 'clear mistake' (on Raz's view, a clear mistake is one that is detectable without having to deliberate about the excluded reasons). Secondly, even if the assignment of probabilities does yield a 'clear mistake', then we can slightly alter the probabilities to yield a structurally similar case where it is not clear. Lastly, in any case, there are good reasons to think that this 'clear mistake' exception is not compatible with Raz's overall view (see: Regan 1990, pp. 20-21). Raz, *The Morality of Freedom*, p. 62; Donald H. Regan, "Reasons, Authority, and the Meaning of 'Obey': Further Thoughts on Raz and Obedience to Law," *Canadian Journal of Law and Jurisprudence* 3 (1990): pp. 20-21.

22. Godwin, *Enquiry Concerning Political Justice and Its Influence on Morals and Happiness*, p. 233.

23. To what extent does this approach depart from orthodox understandings of the Service Conception? This is unclear. However, as it turns out, allowing that the weight of reasons is probability-sensitive does fit with Raz's recent views on normativity. Raz now holds that in cases like the above, we should act in accordance with the balance of probability-weighted reasons: "Given two worthwhile ends, other things being equal the reason to pursue the one more likely to be achieved is the better or stronger reason". Raz, *From Normativity to Responsibility*, p. 116.

likely lead you to better conform with the probability-weighted balance of reasons that apply to you, compared to if you acted on your own assessment of those reasons.

In **Questionable Orders**, this leads to the correct verdict that you should not pre-emptively obey your superior officer. Note, however, that there is a tension in this interpretation of the Normal Justification Thesis: namely, there is a trivial sense in which obeying the authority *is* more likely to lead you to better conform to the probability-weighted reasons that apply to you, since doing so *just is* acting on the probability-weighted balance of your reasons. To act otherwise is *guaranteed* to not lead you to better conform to the probability-weighted balance of your reasons. As such, we can set aside the 'more likely' proviso in the Normal Justification Thesis (Probability-Weighted).

At this point of the analysis, it is helpful to note the structural similarities between a probability-weighted version of the Normal Justification Thesis and orthodox expected value theory. Expected value theory holds that we should maximise probability-weighted average value. The notion of 'value' is purely formal, providing a numerical representation of the relative importance of normative considerations. In this way, values provide a useful index for determining the relative importance of the normative considerations that are relevant to our actions. *Expected* value merely represents the idea that the importance of normative considerations (or 'reasons') is sensitive to probability.[24] The probability-weighted interpretation of the Normal Justification thus closely mirrors expected value reasoning. This suggests the following formulation:

The Normal Justification Thesis (MEV): You should treat an authority's commands pre-emptively if and only if doing so maximises expected value (MEV).

If this formulation is correct – and it seems to be suggested by the foregoing line of analysis – then the Service Conception can avoid problematic cases like **Questionable Orders** by ensuring that it tracks the recommendations of expected value theory. Indeed, to do otherwise would lead it to other problematic cases of involving uncertainty.[25]

24. More precisely, expected value theory assumes that value varies in linear proportion to probability. There are, however, ways of allowing expected value theory to account for attitudes to risk (roughly, by making value vary linearly with probability, but non-linearly with quantity). See: Arrow, "The Theory of Risk Aversion"; Pratt, "Risk Aversion in the Small and in the Large"; Broome, *Weighing Goods*.

25. Violations of expected utility theory are proven to lead to problematic verdicts, for example, in sequential choice cases. See: Katie Steele, "What are the Minimal Requirements of Rational Choice? Arguments from the Sequential-Decision Setting," *Theory and Decision* 68, no. 4 (2010): 463–487. For a classic discussion, see: Donald Davidson, J. C. C. McKinsey, and Patrick Suppes, "Outlines of a Formal Theory of Value, I," *Philosophy of Science* 22, no. 2 (1955): 140–160.

Some might worry, however, that this analysis is unhelpful because it simply confirms an existing objection to the Normal Justification Thesis: namely, that it lacks explanatory power because it is silent about what exactly we have most reason to do.[26] For instance, for all we know, given the Normal Justification Thesis (MEV), we have most reason to obey a dictatorship! It all depends on the underlying axiology. As Scott Hershovitz notes, this apparent emptiness of the Service Conception is unsatisfactory, given that:

> [Raz] presents [the Normal Justification Thesis] as an answer to the question "How can it ever be that one person has a duty to subject one's will and judgement to those of another?" But the normal justification thesis hardly answers that question if it can be satisfied simply because an authority passes some other test for legitimacy.[27]

However, I shall argue that whatever explanatory power the Normal Justification Thesis (MEV) loses through lack of *substance*, it makes up for in *structure*, by reconciling our best theory of legitimate authority with our best theory of rational decision-making under uncertainty. As we shall see, this will allow us to gain valuable insights into how legitimate authorities can guide us in cases of uncertainty, and when obedience is unjustified.

Now, before proceeding, some might worry about the robustness of this identification between the Normal Justification Thesis and a decision-theoretic rule such as Maximise Expected Value. For example, one might wonder: Does the identity hold when we adopt the long-run formulation of the Normal Justification Thesis?[28]

The Normal Justification Thesis (Long-Run): You should treat an authority's commands pre-emptively if and only if doing so would more likely lead you over the long run to better conform with the reasons that independently apply to you, compared to if you acted on your own assessment of those reasons.

As it turns out, including the long-run proviso only strengthens the identification with maximising expected value. This is because, over the long-run, expected value maximisers become overwhelmingly more likely to gain higher average value than non-expected value maximisers.[29] This means that over the long-run, pre-emptive obedience maximises chances of higher average long-run value if and only if doing so maximises expected value.

In the following section, I show how accepting the Normal Justification Thesis (MEV) strengthens the Service Conception, allowing it to draw on other important

26. Scott Hershovitz, "The Role of Authority," *Philosophers' Imprint* 56, no. 7 (2011): 1–19.
27. ibid., p. 5.
28. Raz, *Ethics in the Public Domain*.
29. William Feller, *An Introduction to Probability Theory and its Applications* (New York: Wiley & Sons, 1968); Johanna Thoma, "Risk Aversion and the Long Run," *Ethics*, 2019, 1–31.

resources in decision theory. This will yield a generalised version of the Service Conception of authority that can systematically account for all cases of uncertainty.

5.3 The Value of Legitimate Authority

Having received an order from a legitimate authority, one choice you must make is whether it is better to ignore the order or to obey it. The Service Conception implies that if the authority is legitimate, then it is prospectively better to obey the authority rather than ignore it and follow your own judgement. One source of justification is that the authority is a more reliable judge than you. Another is that pre-emptively obeying its orders is more efficient, saving you the costs of information-gathering and deliberation. In this section, we will model the 'reliability justification' with expected value theory. Drawing on deference principles and I.J. Good's proof of the principle of total evidence will allow us to establish that the expected value of obedience to a legitimate authority whose orders are freely accessible is never prospectively worse than – and is at least sometimes better than – ignoring the order.[30] With this result in hand, it is a trivial exercise to add cost-saving considerations to the model (say, by adding or subtracting values to the possible outcomes, reflecting the costs incurred or saved by deciding for oneself versus obeying the authority).

To illustrate how greater reliability allows an authority to satisfy the Normal Justification Thesis (MEV), consider the following case:

Reliable Order: You are deliberating about whether to commandeer the civilian's van or to use a military-issue truck. If the road ahead is paved, then you should use the van (the truck is slower). If the road is unpaved, you should use the truck (in such conditions, the van would be slower). You think that the road ahead is unpaved. In these cases, you take yourself to be 60% reliable about the correct course of action. You are about to receive an order from your superior officer, who you know to be right in these cases 90% of the time. Your decision situation is shown in Figure 5.1.

In this case, given that the authority's order reveals information about whether the road is likely to be paved or unpaved, choosing to pre-emptively obey has greater expected value, whatever the content of the order turns out to be. This is formally similar to a value of information calculation, showing that free, reliable information never makes a decision situation worse and sometimes makes it better.[31] Informally, the screening-off of your judgement is justified by the fact that including your own

30. Good, "On the Principle of Total Evidence."

31. Good, "On the Principle of Total Evidence"; Bradley and Steele, "Can Free Evidence Be Bad? Value of Information for the Imprecise Probabilist."

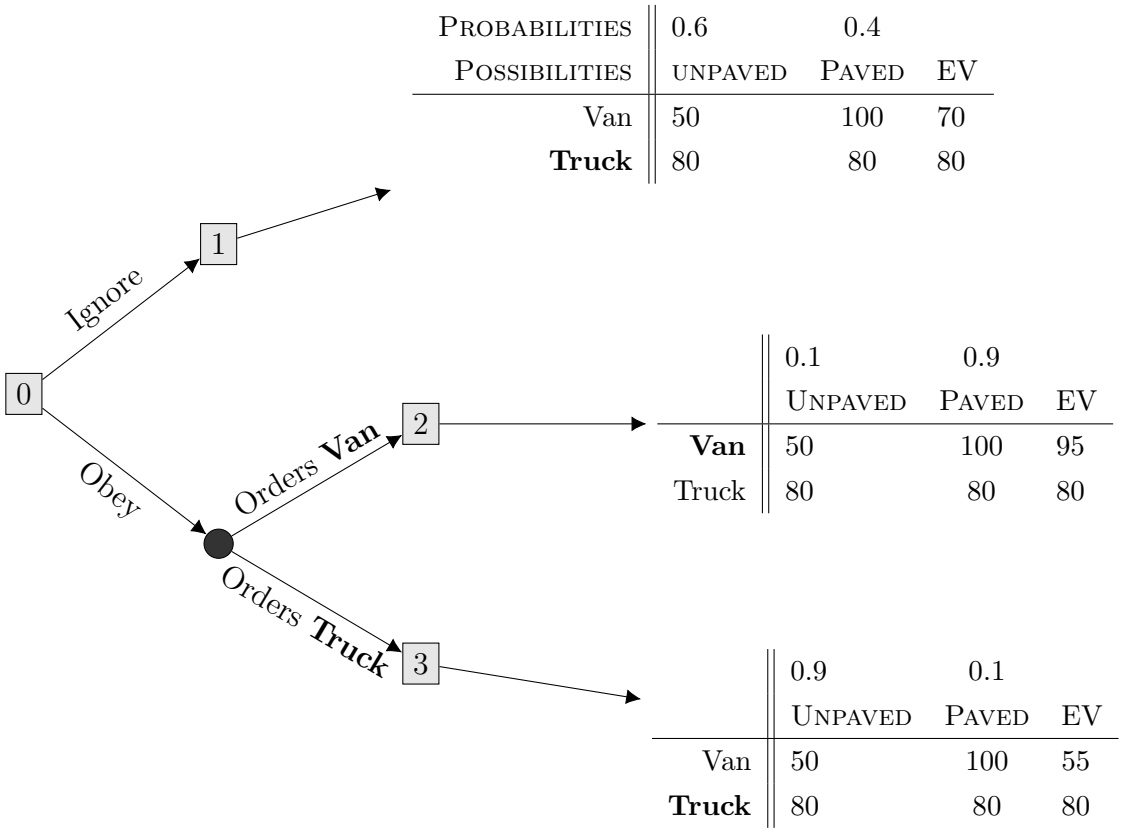


Figure 5.1: Value of Legitimacy

beliefs (say, by combining your prior judgement with the authority's) would amount to 'muddying the waters' or 'introducing noise' into the analysis of the situation, leading you to a prospectively worse set of choices.[32]

This decision-theoretic approach also gives a principled way of determining when an authority should or should not be treated as having legitimate authority. Firstly, we need to be sufficiently sure that the authority is more reliable than us. This turns out to be consistent with Raz's position – which we discussed earlier in relation to the knowability constraint – that reliable beliefs about an authority's legitimacy are generally necessary for improved conformity with reason.[33] On the decision-theoretic analysis, if there is no available information about the authority's legitimacy, then acting on the basis of its orders does not necessarily have determinately greater (or equal) expected value compared to ignoring it. As such, you may be justified in disobeying. For what it is worth, the Anarchist agrees with this verdict:

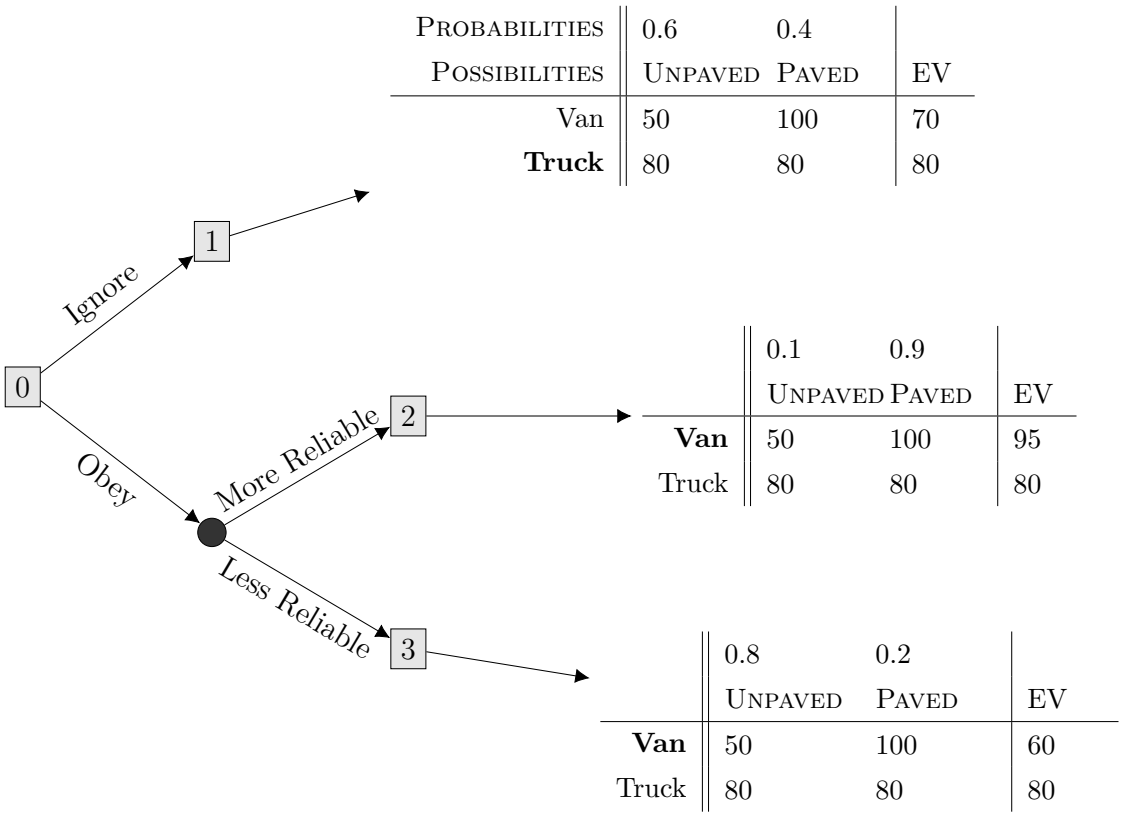
> Wherever I have good reason to believe that another person knows better than myself what is proper to be done, there I ought to conform to his direction.

32. On expert principles in epistemology, see: Richard Pettigrew, "Accuracy, Chance, and the Principal Principle," *Philosophical Review* 121, no. 2 (2012): 241–275; Richard Pettigrew and Michael G Titelbaum, "Deference Done Right," *Philosopher's Imprint* 14, no. 35 (2014): 1–19.

33. Raz, "The Problem of Authority: Revisiting the Service Conception," p. 1025.

> But the advantage which he possesses must be obvious, otherwise I shall not be justified in my proceeding.[34]

To illustrate, suppose now that you are uncertain about whether your authority is more reliable than you. As such, you must weigh up the prospects of following a reliable authority versus an unreliable one (see: Figure 5.2). Suppose that your superior orders you to take the van. If they are reliable, then it is 0.9 probable that they are correct, such that the expected value of obeying is 95. However, if they not more reliable, then (due to their incompetence or malice) it will only be 0.2 probable that they are correct, such that following their order will have an expected value of 60. In this case, you do better by obeying the authority so long as you are sufficiently confident that it is more reliable (in this case, this means being at least about 60% confident about this). Any less and you would be not justified in proceeding.



Probabilities	0.6	0.4	
Possibilities	Unpaved	Paved	EV
Van	50	100	70
Truck	80	80	80

	0.1	0.9	
	Unpaved	Paved	EV
Van	50	100	95
Truck	80	80	80

	0.8	0.2	
	Unpaved	Paved	EV
Van	50	100	60
Truck	80	80	80

Figure 5.2: Uncertain Reliability

A second case in which pre-emptive obedience is not better than ignoring the order is when determining the reliability of the authority is too costly to be worth to potential improvement. This, also, is consistent with Raz's position:

34. Godwin, *Enquiry Concerning Political Justice and Its Influence on Morals and Happiness*, p. 236.

> How much it can be expected to improve our conformity to reason, and how important the matter is, establish what inquiry is reasonable to undertake. When reasonable inquiry will not reveal the case for authority, that case, if it exists at all, is unknowable. It follows that people are not subject to any authority regarding those matters. This argument is used here to establish not merely that it is not rational, or not worthwhile, to carry on with the inquiry about the existence of certain reasons, but that those reasons, authoritative directives, do not exist. There is no authority over the matter, because to exist, authorities must be knowable [in the sense that the cost of acquiring their orders is not greater than the benefit].

In the previous example, if the time it takes you inquire about the authority's reliability costs, say, 15 lives, then the expected value of the inquiry is not worth the cost. This would be another case where you should decide on your own.

Admittedly, these are merely first steps towards reconciling our theories of legitimate authority and rational decision-making under risky situations. The foregoing suggests that modelling the Service Conception in decision-theoretic terms gives a systematic and precise explanation for when pre-emptive obedience is better than deciding alone, and when it is not. Moreover, it seems largely consistent with Raz's earlier stated views on authority and his recent work on reasons. Overall, this analysis sketches how the Service Conception's refutation of Anarchism extends into cases involving uncertainty, by showing when and why pre-emptive obedience is morally and rationally justifiable.

5.4 Objections

A number of objections have been presented to the kind of approach I have sketched above. In particular, Scott Shapiro (2004) argues against any approach that makes the "assumption that willing obedience to authoritative directives is the outcome of some form of decision-making."[35] He presents a litany of objections to decision-theoretic approaches to authority, which he calls the Decision Model. He argues that "not only is the Decision Model false in some type of cases, it is false in all types of cases."[36] Below, I introduce and respond to his objections.

5.4.1 Motivational Weakness

Suppose that you correctly believe that you should take the van, since the road is likely to be paved. However, suppose also that if you were to act on those reasons for taking the van, you would fail due to weakness of will (you always wanted to drive the truck, and you are sure you will somehow delude yourself into believing

35. Shapiro, *Authority*, p. 415.
36. ibid., p. 420.

that you should take it, instead). In this case, Shapiro argues that if you deliberate about whether to obey your superior officer, then you will be unable to benefit from his order, since your weakness of will stops you from being able to follow the order. Shapiro takes this to show that "authoritative directives can serve the benefits they are meant to serve just in case we think that we have the ability to [causally] constrain our future selves."[37]

However, these cases do not undermine the expected value approach given here. The expected value approach determines what you ought to do, given a description of the decision problem. Shapiro's contention is about *how* decision problems should be described: specifically, he argues that we should describe decision problems so that they do not include particular future actions (namely, disobedience) as being available for choice. His contention thus aims at a higher-order decision problem: how should we describe decision problems, given our behavioural tendencies to respond to particular kinds of decision problems? Shapiro suggests that weak-willed agents should treat disobedience as not an available future option. Note, however, that his approach can be justified by adopting an expected value approach to second-order decisions, whereby we assign probabilities to our acting in an optimal fashion, given the choices available to us.[38] If the expected value of treating our future choices as constrained is optimal, then we should adopt Shapiro's approach. However, in other cases, we can treat disobedience as an available, if irrational, option. Shapiro's mistake is that of assuming that just because a simple decision-theoretic approach does not give advice about how to describe decision problems, that it therefore cannot do so.

5.4.2 Guiding Significantly Uncertain Agents

Shapiro argues that the following kinds of cases show that the Decision Model is false:

Radical Cluelessness: You have maximally indeterminate credences regarding what to do. Luckily, an authority who you take to be reliable has just told you what to do. You update your credences, but you are still clueless.

Shapiro holds that this is a counterexample to the Decision Model, since it does not give the correct verdict about what you ought to do: he submits that you should consider yourself bound to follow the authority's order.[39] More generally, he states

37. ibid., pp. 229-230.

38. One might worry here that we cannot rationally assign probabilities to our own choices. For arguments to the contrary, see: Alan Hájek, "Deliberation Welcomes Prediction," *Episteme* 13, no. 4 (2016): 507–528.

39.

that: "[Isaac] Levi has shown that it is usually not possible for a rational agent to harness the informational value of theoretical authorities if the agent were free not to follow the advice, but nevertheless treats the advice as reliable evidence."[40]

There are a number of problems with this line of argument. Firstly, Shapiro's argument appropriates Levi's argument against a specific theory of inductive logic called Objectivist Necessitarianism. The Decision Model is not committed to this theory, or indeed any particular theory of inductive logic. Indeed, even Levi muses that "[p]erhaps no one has ever been a strict objectivist necessitarian."[41] Secondly, Levi's argument does not support the idea that agents should not be free to follow advice from theoretical authorities. It merely supports the idea that any adequate theory of formal epistemology should include the notion of credal commitment, which is a rule governing how we should epistemically respond to particular kinds of evidence. This notion of credal commitment is orthogonal to Shapiro's view that benefiting from authority requires us to act as though we could not possibly disobey. Indeed, if anything, the idea of credal commitment is better suited to the Decision Model.[42]

Shapiro further argues that the Decision Model cannot "validate acceptance of expert advice when the agent's epistemic state is highly indeterminate [in the sense of assigning a low, but precise, credence to a range of possibilities]."[43] He has in mind the following kind of scenario:

Mild Cluelessness: You have mildly indeterminate credences regarding what to do. Luckily, an authority you take to be reliable has just told you what to do. You update your credences, but you are still not sure what to do.

Here, Shapiro is erroneously assuming that accepting a decision-theoretic approach requires endorsing an acceptance-based epistemology, whereby propositions are 'accepted' just in case they are assigned a sufficiently high probability (or likelihood). On the decision-theoretic approach presented earlier, so long as the authority's order improves your later choice situations, then the authority is legitimate and you should obey, even if you do not necessarily 'accept' or 'believe' the proposition implied by their order. However, as we saw earlier, if you are not sufficiently certain about the authority's reliability, then you should not obey the authority – rather, you should decide for yourself.

40. Shapiro, *Authority*, p. 421.
41. Levi, *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*.
42. Shapiro, *Authority*, p. 421; Levi, *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*.
43. Shapiro, *Authority*, p. 423.

5.4.3 Deciding to Decide to Obey

Shapiro also contends that no decision-theoretic approach can account for cases where authorities save us from having to undertake costly and risky deliberation. Consider the following scenario:[44]

Deciding to Obey: Your commanding officer – whose judgement you consider to be as reliable as your own – has just issued you an order. So you think to yourself: should I just obey automatically (saving time and effort) or deliberate for myself?

Shapiro argues that by treating obedience as optional, you have already lost the benefits of authority in this case. Even worse, it seems that you face a (potentially infinite) regress of deciding how to decide. You must deliberate about whether deliberation is most appropriate, which itself is a mode of decision-making that must be justified by further deliberation, and so on.

In the face of this objection, we must ask: what is the alternative course of action? In general, given the risks associated with obeying an illegitimate authority, we must always assess whether obedience to authority is more reliable or efficient than following our own judgement. In this case, you have assessed the authority to be no more reliable than yourself. However, that is not enough to determine its legitimacy. You must also assess whether obedience is more efficient. The authority will be legitimate on this efficiency score only if assessing its legitimacy is worth the potential savings in time and cost. If this assessment is more costly than the potential efficiency gains, then the authority is not legitimate. Pre-emptive obedience is "optimal from an ex ante perspective" only once we can trust that the authority is legitimate.[45] To act other would be reckless and – most likely – irrational.

As for the regress, this is a general problem for theories of practical rationality, for which there are solutions. For instance, we can hold that you should deliberate in a way that is *robustly goal-conducive*, in a sense defined by Hanti Lin (2014).[46] This approach can take into account time and effort costs in deliberations about how to deliberation, while also allowing that even after hearing the order, you can choose to disobey (though, as before, doing so may be irrational). In general, there is no need to treat disobedience as an 'unavailable' option. Even after choosing to obey an authority, it is always open to us to disobey, even if doing so is contrary to reason.

44. ibid., pp. 423-425.
45. ibid., p. 423.
46. Lin, "On the Regress Problem of Deciding How to Decide"

5.4.4 Coordination Problems

Shapiro further argues that "if authorities are able to solve coordination problems, then the Decision Model cannot be correct."[47] He presents the following case:

> The Decision Model claims that it is rational for a player, call him X, to decide to comply with such a directive when, and only when, X can establish that it provides good evidence about other players' behaviour. But the directive provides good evidence about others' behaviour only when it would be rational for others to follow it. However, if these players are rational, the question of whether it is rational for them to follow the directive is the same as whether it is rational for X to follow the directive. Hence, X can establish that the directive constitutes good evidence only if X can first establish that it is rational for him to follow the directive. X has now travelled in a circle. If X wants to establish the rationality of his following the directive, it seems that he must already know that it is rational for him to follow the directive. But since he is trying to establish the rationality of following the directive, he cannot assume the proposition for the purpose of proving it. So, if X does not already believe that it is rational for him to follow the directive, he will never come to that conclusion.[48]

This argument is unsound because it ignores the fact that evidence about the authority and its ability to secure coordination can be sufficient to guide action, even if we do not know directly what is rational for others to do. To illustrate, consider the following case:

Coordination: You are driving to a location to handover the supplies to the local medical staff. However, during your call with them, the phone line went down, preventing you from coordinating on a location. Either they will meet you at the top of the hill (Top) or at the bottom of the hill (Bottom) – you are not sure which. Luckily, you have just received notification from an authority to meet at the bottom of the hill. Your evidence suggests that this authority has a very good track record in securing coordination. Figure 5.3 illustrates the decision situation:

In this case, nothing prevents you from coordinating appropriately with the medical staff on the basis of the authority's directive. Since you consider the authority to be legitimate, you take its order as giving you evidence about what will occur, given that you choose one location or the other. Thus, we have a case where our decision-theoretic approach to authority secures coordination, falsifying Shapiro's claim.

47. Shapiro, *Authority*, p. 424.
48. ibid., pp. 426-427.

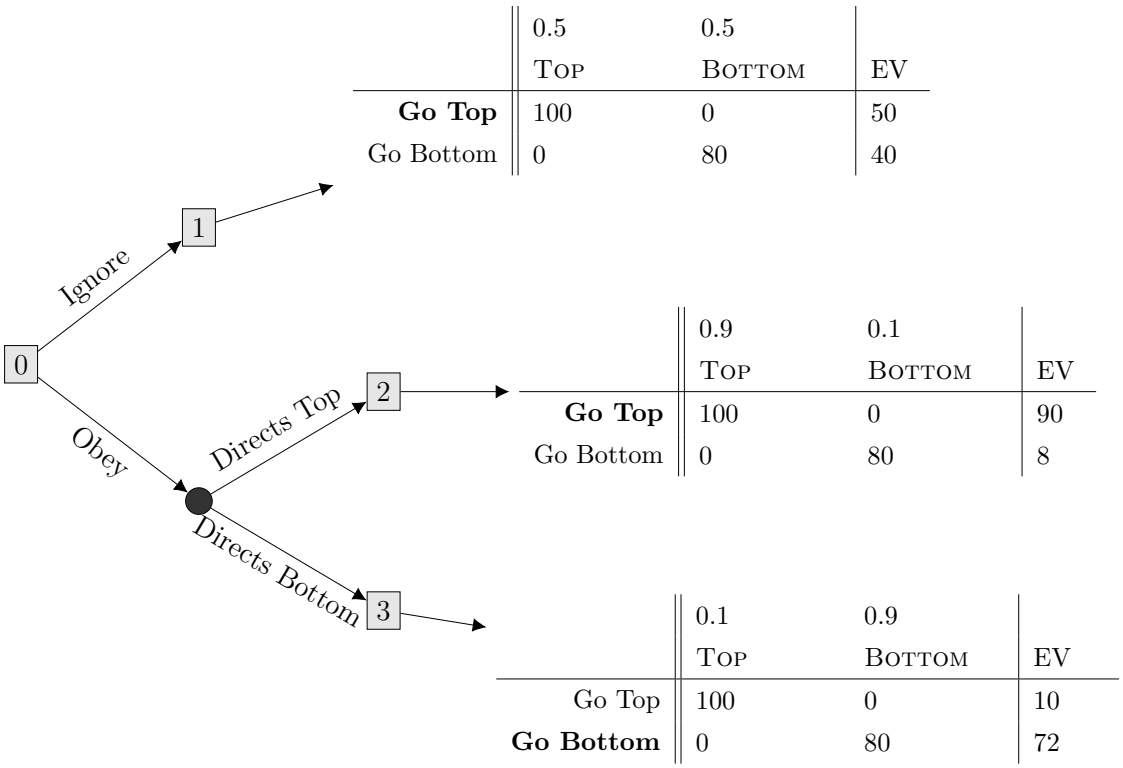


Figure 5.3: How Legitimate Authority Solves Coordination Problems

5.4.5 Practical Versus Theoretical Authority

One final objection to the decision-theoretic approach given here is that it seems to confuse practical authority with theoretical authority. Legitimate *practical* authority, properly understood, is about how agents can give other agents reasons for action by issuing commands. By contrast, theoretical authorities merely give reasons for belief.[49] The objection is that rendering the Service Conception with expected value theory reduces legitimate practical authority to legitimate theoretical authority.

However, this does not follow. If the distinction between practical and theoretical authorities rests on the kinds of speech-acts that they employ – commanding versus advising – then this is consistent with the formal approach given above. Alternatively, if the distinction tracks differing social or institutional roles, a more nuanced decision-theoretic approach could take into account the risks of enshrining, through obedience, an authority who is not duly empowered by the relevant power-conferring social or institutional norms.[50] Just as it is easy to underestimate the flexibility of the Service Conception, it is easy to underestimate the flexibility of the decision-theoretic model of it provided here.

As it stands, none of the examples given above appear to be cases where legiti-

49. Raz, "Legitimate Authority"; Green, *The Authority of the State*, ch. 1.

50. Andrei Marmor, "An Institutional Conception of Authority," *Philosophy & Public Affairs* 39, no. 3 (2011): 238–261; Hershovitz, "The Role of Authority."

mate authorities are merely theoretical authorities. To see why, note that theoretical authorities typically give us first-order evidence or second-order evidence.[51] If they give first-order evidence, we should revise our beliefs on the basis of that evidence and our prior beliefs about what is best to do. This is not what occurs in the decision-theoretic account given above: the authority's order screens-off our understanding of the situation. One might then argue that this sounds more like the authority is giving second-order evidence that undermines our evidence.[52] For example, if our eyes tell us that a room is red, that first-order sensory evidence may be undermined by someone who gives us reason to believe that our seeing red is a drug-induced hallucination. In that case, an appropriate response might be to suspend our judgement about whether the room is red. Clearly, however, this is not a plausible interpretation of the cases given above: the authority's order does not give us reason to suspend our belief in our assessment of the situation. Rather, it gives us reason to do as it says because it is a more reliable or cost-effective guide to the situation than following our own judgement. The expected value approach presented above simply provides a way of measuring how legitimate authority changes our reasons for action in the face of uncertainty.

Of course, there is much more to be said on this point. It may very well be that, in cases of uncertainty, the divide between practical and theoretical authority is not as great as many previously thought. Then again, it may well turn out that the approach given here offers too thin a conception of the nature of legitimate authority. If so, then those who seek a thicker notion of legitimate authority will have to provide an alternative account of how legitimate authority is consistent with rational decision-making in risky cases.

5.5 Conclusion

The Service Conception of authority can – and, indeed, should – be understood in a way that conforms with our best formal theories of rational decision. This approach allows us to precisely determine the conditions under which pre-emptive obedience to authority is rational, given our evidence. Once further developed and defended, this account may complete Raz's answer to the Philosophical Anarchist, showing that even in cases of uncertainty, obedience to authority can be entirely consistent with the decisions of our own understanding, and the dictates of our conscience.

51. Of course, they could give still higher higher-order evidence, but those cases are even less plausible than those discussed here.

52. Thanks here to Sergio Tenenbaum for discussion.

Conclusion

The world is a risky place. Given this fact, what does morality require of us? Many have argued that morality requires us to compromise, to make trade-offs that seem morally unpalatable. In their eyes, anyone who resists compromise – by say, prioritising an innocent individual's life over *any* amount of trivial benefits for others – is a *Moral Absolutist*. These critics argue that absolutist moralities cannot guide us through a risky world because they require us to perform actions that are either absurdly risk-averse, self-effacing, or simply contradictory. As such, these critics conclude that moral compromise is inevitable: lives *can* be outweighed by trivial benefits. I called this line of critique: the Problem of Risk.

In this dissertation, I argued that these critics are mistaken: using an appropriate decision-theoretic approach, almost all absolutist moral theories give adequate verdicts about what to do in a risky world. I argued as follows:

Chapter 1: Prohibition and Probability introduced the debate over the Problem of Risk. I presented a formal feature, called Option Absolutism, that commits a moral theory to the Problem of Risk, and I argued that almost all broadly 'absolutist' theories can reject Option Absolutism in favour of Relational Absolutism.

Chapter 2: Moral Priorities Under Risk gave an example of how one class of so-called absolutist theories – lexical priority theories – can adopt a Relational Absolutist structure using lexicographic decision theory.

Chapter 3: Priorities and Uncertainties gave a more orthodox expected value approach to absolutism, using asymptotic value functions. It pushed the debate towards issues of whether individual or sequences of actions are the proper objects of moral evaluation.

Chapter 4: Duty and Ignorance demonstrated how value of information calculations from decision theory can allow deontologists to correctly determine when to gather more information before acting.

Chapter 5: Authority, Obedience, and Uncertainty argued that the seemingly absolutist Service Conception of authority can be extended, using an expected value approach, to cases where we are uncertain about an authority's legitimacy.

Despite the positive results found in the chapters above, more work is required to determine that the formal decision models presented are the *most* suitable ones available for modelling the moral theories in question. Nevertheless, these chapters demonstrate a number of overlooked points in the discussion of the Problem of Risk. Firstly, they show the versatility of decision-theoretic modelling as a tool for representing moral theories. Secondly, they help us to identify various gaps in our existing moral theorising, including the rationales for positing particular kinds of priority relations between various moral considerations. Thirdly, and most importantly, they show that morality can guide us through a risky world, without compromise.

Bibliography

Aboodi, Ron, Adi Borer, and David Enoch. "Deontology, Individualism, and Uncertainty: A Reply to Jackson and Smith." *Journal of Philosophy* 105, no. 2 (2008): 259–272.

Altham, J. E. J. "Ethics of Risk." *Proceedings of the Aristotelian Society, New Series* 83 (1983): 15–29.

Anscombe, F. J., and R. J. Aumann. "A Definition of Subjective Probability." *The Annals of Mathematical Statistics* 34, no. 1 (1963): 199–205.

Anscombe, G. E. M. *Ethics, Religion and Politics*. Oxford: Basil Blackwell, 1981.

———. "Modern Moral Philosophy." In *Virtue ethics*, edited by Roger Crisp and Michael A Slote, 26–44. Oxford: Oxford University Press, 1997.

———. "War and Murder." In *Moral Problems: A Collection of Philosophical Essays*, edited by James Rachels, 393–407. New York: Harper & Row, 1979.

Arntzenius, Frank, Adam Elga, and John Hawthorne. "Bayesianism, infinite decisions, and binding." *Mind* 113, no. 450 (2004): 251–283.

Arrhenius, Gustaf. "Superiority in Value." *Philosophical Studies* 123, nos. 1-2 (2005): 97–114.

Arrhenius, Gustaf, and Wlodek Rabinowicz. "Millian Superiorities." *Utilitas* 17, no. 02 (2005): 127–146.

———. "Value Superiority." In *The Oxford Handbook of Value Theory*, 225–243. Oxford University Press, 2015.

Arrow, Kenneth J. "The Theory of Risk Aversion." Chap. 2 in *Aspects of the Theory of Risk Bearing*. Helsinki: Yrjo Jahnsonin Saatio, 1965.

Aumann, Robert J. "Utility Theory without the Completeness Axiom." *Econometrica* 30, no. 3 (1962): 445–462.

Bales, R Eugene. "Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?" *American Philosophical Quarterly* 8, no. 3 (1971): 257–265.

Bartha, Paul. "Taking Stock of Infinite Value: Pascal's Wager and Relative Utilities." *Synthese* 154, no. 1 (January 2007): 5–52.

Binmore, Ken, and Alex Voorhoeve. "Defending Transitivity against Zeno's Paradox." *Philosophy and Public Affairs* 76, no. 3 (2003): 272–279.

Bjorndahl, Adam, Alex John London, and Kevin J. S. Zollman. "Kantian Decision Making Under Uncertainty: Dignity, Price, and Consistency." *Philosopher's Imprint* 17, no. 7 (2017): 1–22.

Bradley, Seamus, and Katie Steele. "Can Free Evidence Be Bad? Value of Information for the Imprecise Probabilist." *Philosophy of Science* 83, no. 1 (2016): 1–28.

Briggs, R.A. "Costs of abandoning the Sure-Thing Principle." *Canadian Journal of Philosophy* 45, nos. 5-6 (2015): 827–840.

———. "Normative Theories of Rational Choice: Expected Utility Theory." *The Stanford Encyclopedia of Philosophy*, 2017.

Broome, John. "No Argument against the Continuity of Value: Reply to Dorsey." *Utilitas* 22, no. 04 (November 2010): 494–496.

———. *Weighing Goods*. Cambridge, Massachusetts: Blackwell, 1995.

Brown, Campbell. "Maximalism and the Structure of Acts." *Noûs* 0 (2017): 1–20.

Buchak, Lara. "Decision Theory." In *The Oxford Handbook of Probability and Philosophy*, edited by Alan Hájek and Christopher Hitchcock, 789–814. Oxford: Oxford University Press, 2016.

———. *Risk and Rationality*. Oxford: Oxford University Press, 2013.

Carlson, Erik. "Aggregating Harms — Should We Kill to Avoid Headaches?" *Theoria* 66, no. 3 (2000): 246–255.

———. "Organic Unities, Non-Trade-Off, and the Additivity of Intrinsic Value." *Journal of Ethics* 5 (2001): 335–360.

Chang, Ruth. "Raz on Reasons, Reason, and Rationality: On Raz's From Normativity to Responsibility." *Jerusalem Review of Legal Studies* 8, no. 1 (2013): 199–219.

Colyvan, Mark, Damian Cox, and Katie Steele. "Modelling the Moral Dimension of Decisions." *Noûs* 44, no. 3 (September 2010): 503–529.

Cusbert, John. "Acting on Essentially Comparative Goodness." *Thought: A Journal of Philosophy* 6, no. 2 (2017): 73–83.

Dancy, Jonathan. *Moral Reasons*. xiii, 274 p. Oxford: Blackwell, 1993.

Davidson, Donald, J. C. C. McKinsey, and Patrick Suppes. "Outlines of a Formal Theory of Value, I." *Philosophy of Science* 22, no. 2 (1955): 140–160.

Dietrich, Franz, and Christian List. "What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories." *The Philosophical Review* 126, no. 4 (2017): 421–479.

Dorsey, Dale. "Headaches, Lives and Value." *Utilitas* 21, no. 01 (2009): 36–58.

———. "Headaches, Lives and Value." *Utilitas* 21, no. 1 (2009): 36–58.

Elga, Adam. "Subjective Probabilities should be Sharp." *Philosophers' Imprint* 10, no. 5 (2010): 1–11.

Elliott, Edward. "A Representation Theorem for Frequently Irrational Agents." *Journal of Philosophical Logic* 46, no. 5 (2017): 467–506.

Feller, William. *An Introduction to Probability Theory and its Applications*. New York: Wiley & Sons, 1968.

Finnis, John. *Moral Absolutes*. Catholic University of America Press, 1991.

Fraassen, Bas van. *Laws and Symmetries*. Oxford: Clarendon Press, 1989.

Gewirth, Alan. "Are There Any Absolute Rights?" *The Philosophical Quarterly* 31, no. 122 (1981): 1–16.

Godwin, William. *Enquiry Concerning Political Justice and Its Influence on Morals and Happiness*. Toronto: University of Toronto Press, 1946.

Goldman, Holly S. "Dated Rightness and Moral Imperfection." *The Philosophical Review* 85, no. 4 (1976): 449–487.

Good, I. J. "On the Principle of Total Evidence." *The British Journal for the Philosophy of Science* 17, no. 4 (1967): 319–321.

Graham, Peter A. "In Defense of Objectivism about Moral Obligation." *Ethics* 121, no. 1 (2010): 88–115.

Green, Leslie. *The Authority of the State*. Oxford: Clarendon Press, 1988.

Griffin, James. *Well-Being: Its Meaning, Measurement and Moral Importance*. New York: Oxford University Press, 1986.

Hájek, Alan. "Deliberation Welcomes Prediction." *Episteme* 13, no. 4 (2016): 507–528.

———. "Waging War on Pascal's Wager." *The Philosophical Review* 112, no. 1 (2003): 27–56.

Hájek, Alan, and Christopher Hitchcock, eds. *The Oxford Handbook of Probability and Philosophy*. Oxford: Oxford University Press, 2016.

Hájek, Alan, and Michael Smithson. "Rationality and Indeterminate Probabilities." *Synthese* 187 (2012): 33–48.

Hansson, Sven Ove. "Ethical Criteria of Risk Acceptance." *Erkenntnis* 59, no. 3 (2003): 291–309.

———. *The Ethics of Risk: Ethical Analysis in an Uncertain World*. Palgrave Macmillan, 2013.

Hare, Caspar. "Obligation and Regret When There is No Fact of the Matter About What Would Have Happened if You Had not Done What You Did." *Noûs* 45, no. 1 (2011): 190–206.

———. "Should We Wish Well to All?" *The Philosophical Review* 125, no. 4 (2016): 451–472.

———. *The Limits of Kindness*. Oxford: Oxford University Press, 2013.

———. "Voices from Another World: Must We Respect the Interests of People Who Do Not, and Will Never, Exist?" *Ethics* 117, no. 3 (2007).

Hausner, Melvin. "Multidimensional Utilities." Chap. 12 in *Decision Processes*, edited by R. M. Thrall, C. H. Coombs, and R. L. Davis. New York: John Wiley / Sons, Inc., 1954.

Hawley, Patrick. "Moral Absolutism Defended." *Journal of Philosophy* 105, no. 5 (2008): 273–275.

Hayenhjelm, Madeleine, and Jonathan Wolff. "The Moral Problem of Risk Impositions: A Survey of the Literature." *European Journal of Philosophy* 20, no. 51 (2011): E26–E51.

Hedden, Brian. "Options and Diachronic Tragedy." *Philosophy and Phenomenological Research* 87, no. 1 (June 2013): 1–35.

Hershovitz, Scott. "The Role of Authority." *Philosophers' Imprint* 56, no. 7 (2011): 1–19.

Horty, John. *Reasons as Defaults*. Oxford University Press, 2012.

Houy, Nicolas, and Koichi Tadenuma. "Lexicographic compositions of multiple criteria for decision making." *Journal of Economic Theory* 144, no. 4 (2009): 1770–1782.

Huemer, Michael. "Lexical Priority and the Problem of Risk." *Pacific Philosophical Quarterly* 91 (2010): 332–351.

Isaacs, Yoaav. "Duty and Knowledge." *Philosophical Perspectives* 28, no. 1 (2014): 95–110.

Jackson, Frank. "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101, no. 3 (1991): 461–482.

———. "How Decision Theory Illuminates Assignments of Moral Responsibility." Chap. 2 in *Intention in Law and Philosophy*, edited by Ngaire Naffine, Rosemary J Owens, and John Williams, 19–36. Aldershot, UK: Ashgate, 2001.

Jackson, Frank, and Robert Pargetter. "Oughts, Options, and Actualism." *The Philosophical Review* 95, no. 2 (1986): 233–255.

Jackson, Frank, and Michael Smith. "Absolutist Moral Theories and Uncertainty." *The Journal of Philosophy* 103, no. 6 (2006): 267–283.

Jeffrey, Richard C. *The Logic of Decision*. Second. Chicago and London: The University of Chicago Press, 1983.

Jensen, Kartsen Klint. "Millian Superiorities and the Repugnant Conclusion." *Utilitas* 20, no. 3 (2008): 279–300.

Joyce, James M. "A Defense of Imprecise Credences in Inference and Decision Making." *Philosophical Perspectives* 24 (2010): 281–323.

———. "How Probabilities Reflect Evidence." *Philosophical Perspectives* 19 (2005).

———. *The Foundations of Causal Decision Theory*. xii, 268 p. Cambridge; New York: Cambridge University Press, 1999.

Kagan, Shelly. "The Additive Fallacy." *Ethics* 99, no. 1 (1988): 5.

Kagan, Shelly. *The Limits of Morality*. Oxford: Clarendon Press, 1991.

Kahneman, Daniel. *Thinking, Fast and Slow*. New York: Farrar, Straus / Giroux, 2011.

Kamm, Frances M. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. Cary, NC: Oxford University Press, 2006.

———. "Supererogation and Obligation." *Journal of Philosophy* 82, no. 3 (1985): 118–138.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Edited by Allen W. Wood. New Haven; London: Yale University Press, 2002.

———. *The Metaphysics of Morals*. Edited by Mary J Gregor (trans.) Cambridge University Press, 1996.

Kaplan, Mark. *Decision Theory as Philosophy*. Cambridge: Cambridge University Press, 1996.

Keynes, John Maynard. *A Treatise of Probability*. London; New York: Macmillan; AMS.

Kirkpatrick, James R. "Permissibility and the Aggregation of Risks." *Utilitas* 30, no. 1 (2017): 107–119.

Kyburg, H.E. *Probability and the Logic of Rational Belief*. Middletown, CT: Wesleyan University Press, 1961.

Lazar, Seth. "Anton's Game: Deontological Decision Theory for an Iterated Decision." *Utilitas* 29, no. 1 (2017): 88–109.

———. "Deontological Decision Theory and Agent-Centered Options." *Ethics* 127, no. 3 (2017): 579–609.

———. "Deontological Decision Theory and the Grounds of Subjective Permissibility." 2018.

———. "In Dubious Battle: Uncertainty and the Ethics of Killing." *Philosophical Studies*, 2017, 1–27.

———. "Limited Aggregation and Risk." 2018.

———. "Risky Killing and the Ethics of War." *Ethics* 126, no. 1 (2015): 91–117.

———. "Risky Killing: How Risks Worsen Violations of Objective Rights." *Journal of Moral Philosophy*, 2017, 1–30.

Lazar, Seth, and Chad Lee-Stronach. "Axiological Absolutism and Risk." *Noûs*, 2017, 1–17.

Lee-Stronach, Chad. "Moral Priorities Under Risk." *Canadian Journal of Philosophy*, 2017, 1–19.

Leitgeb, Hannes. "The Stability Theory of Belief." *Philosophical Review* 123, no. 2 (2014): 131–171.

Levi, Isaac. *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*. Cambridge, Massachusetts; London, England: MIT Press, 1980.

———. "Why Indeterminate Probability is Rational." *Journal of Applied Logic* 7, no. 4 (December 2009): 364–376.

Lin, Hanti. "On the Regress Problem of Deciding How to Decide." *Synthese* 191, no. 4 (2014): 661–670.

Lin, Hanti, and Kevin T. Kelly. "A geo-logical solution to the lottery paradox, with applications to conditional logic." *Synthese* 186 (2012): 531–575.

———. "Propositional Reasoning that Tracks Probabilistic Reasoning." *Journal of Philosophical Logic* 41, no. 6 (2012): 957–981.

Lord, Errol, and Barry Maguire, eds. "The Implementation Problem for Deontology." Chap. 14 in *Weighing Reasons*, 338–354. Oxford: Oxford University Press, 2016.

Luce, R. Duncan, and Howard Raiffa. *Games and Decisions: Introduction and Critical Survey*. New York: John Wiley / Sons, Ltd., 1957.

MacAskill, William. "Normative uncertainty as a voting problem." *Mind* 125, no. 500 (2016): 967–1004.

Marmor, Andrei. "An Institutional Conception of Authority." *Philosophy & Public Affairs* 39, no. 3 (2011): 238–261.

McCarthy, David. "Rights, Explanation, and Risks." *Ethics* 107, no. 2 (1997): 205–225.

McKerlie, Dennis. "Rights and Risk." *Canadian Journal of Philosophy* 16, no. 2 (1986): 239–251.

Nagel, Thomas. *Mortal Questions*. Cambridge University Press, 1979.

Norcross, Alastair. "Comparing Harms: Headaches and Human Lives." *Philosophy and Public Affairs* 26, no. 2 (1997): 135–167.

———. "Great harms from small benefits grow: how death can be outweighed by headaches." *Analysis*, no. April (1998): 152–158.

———. "Two Dogmas of Deontology: Aggregation, Rights, and the Separateness of Persons." *Social Philosophy and Policy* 26, no. 01 (2009): 76.

Nover, Harris, and Alan Hájek. "Vexing Expectations." *Mind* 113, no. 450 (2004): 237–249.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic, 1974.

———. "Moral Complications and Moral Structures." *Natural Law Forum* 13 (1968): 1–50.

Oddie, Graham, and Peter Milne. "Act and Value: Expectation and the Representability of Moral Theories." *Theoria*, 1991.

O'Neill, Onora. "Instituting Principles: Between Duty and Action." *Southern Journal of Philosophy* 36, no. S1 (1998): 79–96.

Osiel, Mark J. "Obeying Orders: Atrocity, Military Discipline, and the Law of War." *California Law Review* 86, no. 5 (1998): 943.

Perry, Stephen. "Risk, Harm, Interests, and Rights." In *Risk: Philosophical Perspectives*, edited by Tim Lewens. Routledge: Oxford University Press, 2007.

———. "Second-Order Reasons, Uncertainty and Legal Theory." *California Law Review* 62 (1988): 913–994.

Pettigrew, Richard. "Accuracy, Chance, and the Principal Principle." *Philosophical Review* 121, no. 2 (2012): 241–275.

———. "Accuracy, Risk, and the Principle of Indifference." *Philosophy and Phenomenological Research* 92, no. 1 (2016): 35–59.

Pettigrew, Richard, and Michael G Titelbaum. "Deference Done Right." *Philosopher's Imprint* 14, no. 35 (2014): 1–19.

Portmore, Douglas. "Maximalism versus Omnism about Permissibility." *Pacific Philosophical Quarterly* 98, no. 2017 (2017): 427–452.

Portmore, Douglas W. *Commonsense Consequentialism: Wherein Morality Meets Rationality*. xx, 266 p. Oxford: Oxford University Press, 2011.

Pratt, John W. "Risk Aversion in the Small and in the Large." Chap. Chapter 19 in *Handbook of the Fundamentals of Financial Decision Making*, 317–331. 1964.

Pummer, Theron. "Whether and Where to Give." *Philosophy and Public Affairs* 44, no. 1 (2016): 77–95.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

———. *A Theory of Justice: Revised Edition*. Cambridge, MA: Belknap Press, 1999.

———. *Justice as Fairness: A Restatement*. Edited by Erin Kelly. Cambridge, Massachusetts and London, England: Harvard University Press, 2001.

———. "Some Reasons for the Maximin Criterion." *The American Economic Review* 64, no. 2 (1974): 141–146.

Raz, Joseph. "Authority and Justification." *Philosophy and Public Affairs* 14, no. 1 (1985): 3–29.

———. *Ethics in the Public Domain*. Oxford: Clarendon Press, 1994.

———. "Facing Up: A Reply." *California Law Review* 62 (1988).

———. *From Normativity to Responsibility*. Oxford; New York: Oxford University Press, 2011.

———. "Legitimate Authority." In *Philosophical Law*, edited by Richard Bronaugh, 3–27. Westport, Conn.: Greenwood Press, 1978.

———. *Practical Reason and Norms*. Oxford: Oxford University Press, 1999.

———. *The Morality of Freedom*. Oxford: Clarendon Press, 1986.

———. "The Problem of Authority: Revisiting the Service Conception." *Minnesota Law Review* 90, no. 1979 (2006): 1003–1044.

Regan, Donald H. "Reasons, Authority, and the Meaning of 'Obey': Further Thoughts on Raz and Obedience to Law." *Canadian Journal of Law and Jurisprudence* 3 (1990): 3–28.

Resnik, Michael D. *Choices: An Introduction to Decision Theory*. London: University of Minnesota Press, 1987.

Ridge, Michael. "How to Avoid Being Driven to Consequentialism: A Comment on Norcross." *Philosophy & Public Affairs* 27, no. 1 (1998): 50–58.

Ross, Jacob, and Mark Schroeder. "Belief, Credence, and Pragmatic Encroachment." *Philosophy and Phenomenological ...* 88, no. 2 (2012): 259–288.

Savage, Leonard J. *The Foundations of Statistics*. 2nd rev. New York: Dover, 1972.

Scanlon, T.M. *Moral Dimensions*. Cambridge Massachusetts, and London, England: Belknap Press of Harvard University Press, 2008.

———. *What We Owe to Each Other*. Cambridge, Massachusetts and London, England: The Belknap Press of Harvard University Press, 1998.

Sciaraffa, Stefan. "On Content-Independent Reasons: It's Not in the Name." *Law and Philosophy* 28, no. 3 (2008): 233–260.

Sepielli, Andrew. "What to Do When You Don't Know What to Do." In *Oxford Studies in Metaethics*. 2009.

Shapiro, Scott J. *Authority*, edited by Jules L. Coleman, Kenneth Einar Himma, and Scott J. Shapiro, 382–439. Oxford University Press, 2004.

Shiffrin, Seana Valentine. "Wrongful Life, Procreative Responsibility, and the Significance of Harm." *Legal Theory* 5, no. 2 (1999): 117–148.

Smith, Holly M. "A Paradox of Promising." *Philosophical Review* 106, no. 2 (1997): 153–196.

———. "Making Moral Decisions." *Noûs* 22, no. 1 (1988): 89–108.

———. "The Subjective Moral Duty to Inform Oneself before Acting." *Ethics* 125, no. 1 (2014): 11–38.

Steele, Katie. "What are the Minimal Requirements of Rational Choice? Arguments from the Sequential-Decision Setting." *Theory and Decision* 68, no. 4 (2010): 463–487.

Swenson, Philip. "Subjective Deontology and the Duty to Gather Information." *Ethics* 125, no. October (2016): 257–271.

Taurek, John M. "Should the Numbers Count?" *Philosophy and Public Affairs* 6, no. 4 (1977): 293–316.

Temkin, Larry S. "A Continuum Argument for Intransitivity." *Philosophy and Public Affairs* 25, no. 3 (1996): 175–210.

———. *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford; New York: Oxford University Press, 2012.

Tenenbaum, Sergio. "Action, Deontology, and Risk: Against the Multiplicative Model." *Ethics* 127 (2017): 1–36.

Tessman, Lisa. *Moral Failure: On the Impossible Demands of Morality*. Oxford: Oxford University Press, 2015.

Thoma, Johanna. "Risk Aversion and the Long Run." *Ethics*, 2019, 1–31.

Thomson, Judith Jarvis. *Rights, Restitution, and Risk: Essays in Moral Theory*. Cambridge, Mass.: Harvard University Press, 1986.

———. *The Realm of Rights*. Harvard University Press, 1990.

Tversky, Amos, and Daniel Kahneman. "Judgement Under Uncertainty: Heuristics and Biases." *Science* 185, no. 4157 (1974): 1124–1131.

Vallinder, Aron. "Imprecise Bayesianism and Global Belief Inertia." *British Journal for the Philosophy of Science* 0, no. June (2018): 1–26.

Von Neumann, John, and Oskar Morgenstern. *Theory of Games and Economic Behavior*. 3rd. 1947. Princeton: Princeton University Press, 1990.

Voorhoeve, Alex. "How Should We Aggregate Competing Claims?" *Ethics* 125, no. 1 (2014): 64–87.

Wilford, John Noble. *The Mapmakers*. Revised Edition. New York: Vintage Books, 2000.

Williamson, Timothy. *Knowledge and Its Limits*. Oxford Scholarship Online, 2003.

Wolff, Robert Paul. *In Defense of Anarchism*. University of California Press, 1970.