CANADIAN JOURNAL OF PHILOSOPHY
2018, VOL. 48, NO. 5, 793–811
https://doi.org/10.1080/00455091.2017.1415104

Routledge
Taylor & Francis Group

Moral priorities under risk

Chad Lee-Stronach

School of Philosophy, Australian National University, Canberra, Australia

ABSTRACT

Many moral theories are committed to the idea that some kinds of moral considerations should be respected, whatever the cost to 'lesser' types of considerations. A person's life, for instance, should not be sacrificed for the trivial pleasures of others, no matter how many would benefit. However, according to the decision-theoretic critique of lexical priority theories, accepting lexical priorities inevitably leads us to make unacceptable decisions in risky situations. It seems that to operate in a risky world, we must reject lexical priorities altogether. This paper argues that lexical priority theories can, in fact, offer satisfactory guidance in risky situations. It does so by equipping lexical priority theories with overlooked resources from decision theory.

ARTICLE HISTORY Received 20 March 2017; Accepted 5 December 2017

KEYWORDS Ethics; lexical priority; decision theory; risk; ethics of artificial intelligence; public policy

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

CONTACT Chad Lee-Stronach chad.lee-stronach@anu.edu.au

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. 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 (Huemer 2010; Jackson and Smith 2006, 2016). This is because lexical priority theories appear to commit a decision-maker to at least one of the following problems: the Permissiveness Problem, the 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 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 paper 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.

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

Part 2 briefly sets out the preliminaries for modelling a moral theory using decision theory. **Part 3** responds to the Permissiveness Problem. **Part 4** responds the Risk Problem. **Part 5** responds to the Agglomeration Problem. **Part 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. 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*.[2] 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[3]; 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.[4] 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.[5] 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 considerations are *infinitely* more important than lower-ranked considerations, and so should be represented by an infinite value difference.[6] 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

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

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 (Temkin 2012, chap. 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.

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 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.[7]

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

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

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? (Table 1).

As shown in Table 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.

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 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.[8] One example is lexicographic expected value theory. Unlike most other approaches, the lexicographic approach constitutes a minimal departure from orthodox expected value theory.[9] 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.[10] 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).[11] 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

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

Table 1. The Permissiveness Problem.

	Careless	Careful	Expected value
Increase speed limit	Large increase in pedestrian deaths, but improved passenger satisfaction	No increase in pedestrian deaths and improved passenger satisfaction	Increasing the speed limit has negatively infinite expected value
	$-\infty$	f_1	$-\infty$
	Probability of Careless, given **Increase Speed Limit** = p	Probability of Careful, given **Increase Speed Limit** = (1 − p)	
Maintain speed limit	Slight increase in pedestrian deaths, and no improvement in passenger satisfaction	No increase in pedestrian deaths and no improvement in passenger satisfaction	Maintaining the speed limit has negatively infinite expected value
	$-\infty$	f_2	$-\infty$
	Probability of Careless, given **Maintain Speed Limit** = q	Probability of Careful, given **Maintain Speed Limit** = (1 − q)	

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

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 (since there are no other considerations in this example).[13]

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.

As shown in Table 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.

4. The Risk Problem

Although lexicographic decision theory solves the Permissiveness Problem, it seems to expose lexical priority theories to the 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.[14]

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

As shown in Table 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

Table 2. Lexicographic expected value.

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

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

Table 3. The Risk Problem.

	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

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 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* (Jackson and Smith 2006, 276; Isaacs 2014, 97). 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.[15] 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

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

expected moral value of passenger satisfaction acting as a tie-breaker between equally choice-worthy threshold values.[16]

Note that lexical priority theories can adopt this approach because they are not, strictly speaking, *absolutist* moral theories. Absolutist theories categorically hold that some kinds of actions are always morally prohibited, irrespective of what other options are available to the decision-maker (Jackson and Smith 2006, 268). Since prohibition does not seem to come in degrees or to be amenable to aggregation, 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 (Lin and Kelly 2012a, 2012b). Where standard probability threshold approaches assess whether 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 Risk Problem (Table 3), 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 4).

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

Table 4. The Risk Problem avoided.

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

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

that she increases the speed limit), the official maximises expected moral value within the revised decision problem by increasing the speed limit.[17] In this way, the official upholds her commitment to lexical priorities while also avoiding the Risk Problem. However, as we shall see, positing any kind of probabilistic threshold seems to raise a further problem for lexical priority theories.

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 Risk Problem. Below, I will explain how lexical priority theories can avoid both versions of the Agglomeration Problem, without abandoning a probabilistic threshold.[18]

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 5).[19]

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

Table 5. The Conjunctive Agglomeration Problem.

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	$r < t$

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, there can be cases where the relevant conjuncts fall on one side of the threshold, but the conjunction falls on the other.[20] 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 follows (Table 6).

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 from her deliberations, the official updates her beliefs and maximises expected moral value (Table 7).[21]

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.

Table 6. A more accurate specification of the states.

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[1]	$s = 1 - p - q - r$

[1]The value of s in this table is implied by the description of the decision situation.

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

Table 7. The Conjunctive Agglomeration Problem avoided.

	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

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 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:

(S_1) The public is always careful.
(S_2) The public is careful on weekdays but careless on weekends.

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

(S_3) The public is careless on weekdays but careful on weekends.
(S_4) 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:

(S_{1*}) The public is always careful.
(S_{2*}) 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 S_{1*} or S_{2*} 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.

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

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.[22] Once the decision problem is specified, the official should maximise expected lexicographic moral value.

6. Further uncertainties for lexical priority theories

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 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 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

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

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.[23] 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.[24] 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 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.[25] 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?[26]

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.[27] 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.

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

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.

Notes

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), pp. 37–39. Rawls himself takes lexical priorities to only be a 'useful approximation' for the purposes of his theory.
2. For an accessible introduction to the normative applications of expected utility theory, see: Briggs (2017).
3. 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: Joyce (1999, chap. 5).
4. 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.
5. Slightly more formally, for all considerations C_1 and C_2, a moral importance relation w (where $>_w$ stands for 'more important than' and $\sim_w$ stands for 'equally as important as'), and a moral value function v: $C_1 >_w C_2$ if and only if $v(C_1) > v(C_2)$; and $C_1 \sim_w C_2$ if and only if $v(C_1) = v(C_2)$.
6. See, for example: Jackson and Smith (2006, 2016), Colyvan, Cox, and Steele (2010), Huemer (2010), Hayenhjelm and Wolff (2012), Hansson (2013) and Bjorndahl, London, and Zollman (2017). Huemer's model of lexical priority theories is an informal rendition of the infinite-values model introduced in Jackson and Smith (2006).
7. For versions of this problem, see: Huemer (2010, 337), Bjorndahl, London, and Zollman (2017, 8), Hansson (2013, chap. 2) and Hayenhjelm and Wolff (2012, E26–51).
8. For a review of such approaches, see: Hájek (2003).
9. The Permissiveness Problem can also be avoided without abandoning infinite values, using Relative Utility Theory – see: Bartha (2007). For simplicity of exposition, I have chosen an alternative, lexicographic approach.
10. See, for example: Houy and Tadenuma (2009).
11. Rawls's theory, of course, adopts a lexicographic framework. However, his approach rejects expected utility maximisation altogether, in favour of a maximin decision rule for decision-making under ignorance (Rawls 1974). An early exposition of the idea of lexicographic utility functions is found in Von

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

and Morgenstern (1947). However, it wasn't until later that lexicographic utility functions received their first systematic study in Hausner (1954).

13. Slightly more precisely, for two options $\underline{a} = (a_1, a_2)$ and $\underline{b} = (b_1, b_2)$: $\underline{a} \sim \underline{b}$ if and only if $((a_1 = b_1) \& (a_2 = b_2))$; $\underline{a} > \underline{b}$ if and only if $(a_1 > b_1)$ or $((a_1 = b_1) \& (a_2 > b_2))$. This generalises to *n*-components.
14. See: Jackson and Smith (2006, 2016) and Huemer (2010).
15. This justification is similar to the justification of a probability threshold in: Leitgeb (2014, 150–151). Note that the justification put forward here for ignoring low probabilities differs from that proposed in Smith (2014), in two main respects. Firstly, where Smith posits that practical norms are 'tolerant' of slight deviations from infinite precision, the approach put forward 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: Schroeder and Ross (2012). Secondly, Smith's account merely permits, but does not require, ignoring probabilities below the threshold. The justification offered in this paper is that we are morally required to ignore sufficiently improbable possibilities, since doing so best ensures that we conform to our moral requirements.
16. 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: Lin (2014).
17. This approach is similar to that proposed in: Hawley (2008). One 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.
18. This discussion addresses synchronic versions of the Agglomeration Problem. For a discussion of strategies for solving the diachronic versions of the Agglomeration Problem, see: Lazar and Lee-Stronach (2017).
19. As noted in: Jackson and Smith (2006, 276–278) and Huemer (2010, 336–339).
20. See, for example, Henry Kyburg's epistemological puzzle, the Lottery Paradox, which is structurally almost identical to the Agglomeration Problem. See: Kyburg (1961).
21. 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 (2012a, 2012b).
22. This is one of the key results in: Lin and Kelly (2012a, 2012b).
23. For discussion of this idea, see: Carlson (2001). As applied to ethics and risk, see: Lazar and Lee-Stronach (2017).
24. This idea has been explored with respect to 'the good' by: Carlson (2000) and Broome (2010). For an approach that applies to multiple kinds of values, see: Temkin (2012, chap. 10).
25. A related notion of conditionality of moral worth can be found in, for example: Kant (2002, sec. 1).
26. As far as I am aware, the closest work that systematically addresses this question is: Horty (2012). However, Horty's approach does not seem to be sufficiently general for the purposes of most lexical priority theories, since it eschews both probabilistic uncertainty and the idea that moral considerations can be meaningfully weighed against each other.
27. See: Raz (1999, 178–199). Thanks to Andrew Williams for discussion.

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

Acknowledgements

I am extremely grateful to the generous and vibrant philosophical community at the Australian National University for many helpful discussions and comments on this paper. Thanks in particular to Christian Barry, Paul Bartha, Justin Bruner, Edward Elliot, Frank Jackson, Avery Kolers, Carina Lee-Stronach, Leon Leontyev, Hanti Lin, Daniel Nolan, Alex Sandgren, Lachlan Umbers, David Wiens, Timothy L. Williamson, Shang Long Yeo, various anonymous referees, editors at the *Canadian Journal of Philosophy*, and an audience at the ANU Philsoc Seminar. Thanks also to Gabriel Watts and an audience at the University of Sydney Postgraduate Philosophy Seminar. I owe extra special thanks to Alan Hájek and Seth Lazar for their support and constructive advice on this paper. Any mistakes in this paper are my own.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Chad Lee-Stronach is a PhD candidate at the School of Philosophy, Australian National University. His research lies at the intersection of moral philosophy, decision theory and social choice theory. His PhD dissertation identifies and explores the general conditions that moral theories must meet in order to operate satisfactorily in contexts of risk and uncertainty. Other work related to his dissertation, titled 'Axiological Absolutism and Risk' (co-authored with Seth Lazar), is forthcoming in Noûs.

ORCID

Chad Lee-Stronach http://orcid.org/0000-0003-4351-2404

References

Bartha, Paul. 2007. "Taking Stock of Infinite Value: Pascal's Wager and Relative Utilities." *Synthese* 154 (1): 5–52.

Bjorndahl, Adam, Alex John London, and Kevin J. S. Zollman. 2017. "Kantian Decision Making Under Uncertainty: Dignity, Price, and Consistency." *Philosopher's Imprint* 17 (7): 1–23.

Briggs, R. A. 2017. "Normative Theories of Rational Choice: Expected Utility Theory." *The Stanford Encyclopedia of Philosophy*. Spring 2017 Edition. Edited by Edward N. Zalta. https://plato.stanford.edu/archives/spr2017/entries/rationality-normative-utility/.

Broome, John. 2010. "No Argument Against the Continuity of Value: Reply to Dorsey." *Utilitas* 22 (4): 494–496.

Carlson, Erik. 2000. "Aggregating Harms – Should We Kill to Avoid Headaches?" *Theoria* 66 (3): 246–255.

Carlson, Erik. 2001. "Organic Unities, Non-Trade-Off, and the Additivity of Intrinsic Value." *Journal of Ethics* 5 (4): 335–360.

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press

Colyvan, Mark, Damian Cox, and Katie Steele. 2010. "Modelling the Moral Dimension of Decisions." *Noûs* 44 (3): 503–529.
Hájek, Alan. 2003. "Waging War on Pascal's Wager." *The Philosophical Review* 112 (1): 27–56.
Hansson, Sven Ove. 2013. *The Ethics of Risk: Ethical Analysis in an Uncertain World*. New York: Palgrave Macmillan.
Hausner, Melvin. 1954. "Multidimensional Utilities." In *Decision Processes*, edited by R. M. Thrall, C. H. Coombs and R. L. Davis, 167–180. New York: John Wiley and Sons Inc.
Hawley, Patrick. 2008. "Moral Absolutism Defended." *Journal of Philosophy* 105 (5): 273–275.
Hayenhjelm, Madeleine, and Jonathan Wolff. 2012. "The Moral Problem of Risk Impositions: A Survey." *European Journal of Philosophy* 20 (S1): E26–E51.
Horty, John. 2012. *Reasons as Defaults*. New York: Oxford University Press.
Houy, Nicolas, and Koichi Tadenuma. 2009. "Lexicographic Compositions of Multiple Criteria for Decision Making." *Journal of Economic Theory* 144 (4): 1770–1782.
Huemer, Michael. 2010. "Lexical Priority and the Problem of Risk." *Pacific Philosophical Quarterly* 91 (3): 332–351.
Isaacs, Yoaav. 2014. "Duty and Knowledge." *Philosophical Perspectives* 28 (1): 1–16.
Jackson, Frank, and Michael Smith. 2006. "Absolutist Moral Theories and Uncertainty." *The Journal of Philosophy* 103 (6): 267–283.
Jackson, Frank, and Michael Smith. 2016. "The Implementation Problem for Deontology." In *Weighing Reasons*, edited by Errol Lord and Barry Maguire, 338–354. Oxford: Oxford University Press.
Joyce, James M. 1999. *The Foundations of Causal Decision Theory*. Cambridge; New York: Cambridge University Press.
Kant, Immanuel. 2002. *Groundwork for the Metaphysics of Morals*. Edited by Allen W. Wood. New Haven; London: Yale University Press.
Kyburg, H. E. 1961. *Probability and the Logic of Rational Belief*. Middletown, CT: Wesleyan University Press.
Lazar, Seth, and Chad Lee-Stronach. 2017. "Axiological Absolutism and Risk." *Noûs*: 1–17. doi:10.1111/nous.12210.
Leitgeb, Hannes. 2014. "The Stability Theory of Belief." *Philosophical Review* 123 (2): 131–171.
Lin, Hanti. 2014. "On the Regress Problem of Deciding How to Decide." *Synthese* 191 (4): 661–670.
Lin, Hanti, and Kevin T. Kelly. 2012a. "A Geo-Logical Solution to the Lottery Paradox, With Applications to Conditional Logic." *Synthese* 186: 531–575.
Lin, Hanti, and Kevin T. Kelly. 2012b. "Propositional Reasoning that Tracks Probabilistic Reasoning." *Journal of Philosophical Logic* 41 (6): 957–981.
Neumann, J. Von, and O. Morgenstern. 1947. *Theory of Games and Economic Behaviour*. 2nd ed. Princeton: Princeton University Press.
Rawls, John. 1974. "Some Reasons for the Maximin Criterion." *The American Economic Review* 64 (2): 141–146.
Raz, Joseph. 1999. *Practical Reason and Norms*. Oxford: Oxford University Press.
Schroeder, Mark, and Jacob Ross. 2012. "Belief, Credence, and Pragmatic Encroachment." *Philosophy and Phenomenological Research* 88 (2): 1–30.
Smith, Nicholas J. J. 2014. "Is Evaluative Compositionality a Requirement of Rationality?" *Mind* 123 (490): 457–502.
Temkin, Larry. 2012. *Rethinking the Good*. Oxford; New York: Oxford University Press.

https://doi.org/10.1080/00455091.2017.1415104 Published online by Cambridge University Press