Absolutism and its Limits

John Hawthorne
Yoaav Isaacs
Clayton Littlejohn

Abstract

Many philosophers think that given the choice between saving the life of an innocent person and averting any number of minor ailments or inconveniences, it would be better to save the life. How, then, should one compare the risk of an innocent person's life to such minor ailments and inconveniences? If lives are infinitely more important than insignificant factors then any risk cannot be outweighed, and that's untenable. An alternative approach seems more promising: let the values of such insignificant factors be bounded, as then there will be well-behaved tradeoffs between insignificant things and risk to an innocent life. We argue, however, that bounding the values of insignificant factors poses myriad problems.

Introduction

Many philosophers think that given the choice between saving the life of an innocent person and averting many minor ailments or inconveniences, it would be better to save the life. Similarly, when confronted with the choice between securing many minor or trifling goods while violating someone's rights or respecting this person's rights and missing out on these many minor goods, it would be better to forego these minor goods. Moreover, the specific numbers aren't thought to matter that much. When someone thinks we shouldn't trade a life to cure millions of barely noticeable headaches or to secure millions of cups of coffee, they don't think that haggling about quantity is going to change things. Billions and trillions of headaches / cups of coffee will still be insignificant when compared to a life.[1]

These intuitions concern cases where stakes are certain – X many headaches vs. a life. It is less clear how to accommodate cases with uncertain stakes – X many headaches vs. some nonzero probability of risk to a life. Even taking for granted the intuitions about cases where stakes are certain, there's latitude about cases where stakes are uncertain.

Suppose we posit that the value of a headache or a cup of coffee is constant. That is, each headache or cup of coffee makes the same difference to the overall value of a world however many headaches, cups of coffee, or other things there are in the world; adding a headache to a world always makes it worse by some particular amount, and adding a cup of coffee to a world always makes it better by some particular amount. In this case, it follows that the value of a life must be infinite relative to the value of a headache or a cup of coffee; the value of a life must be greater than any multiple of the value of a headache or a cup of coffee. Note that this is not to say that the value of a life must be infinity—the extended real number, ∞. The extended reals pose all sorts of unpalatable problems for these purposes. For example, $2 \cdot \infty = \infty$, so the value of saving two lives would be no different than the value of saving one. It's far more natural to instead take the values to

[1] See Norcross (1997) and Scanlon (1998) and Dorsey (2009) for expressions of this sensibility.

have a lexical order.[2] That is, if w_1 does better than w_2 in terms of lives saved then w_1 is guaranteed to be better than w_2, and the values of headaches / coffee cups function as a tie-breaker. But this option has some unpalatable consequences regarding uncertain stakes. The view seems to lead to a kind of practical paralysis. Even if one should not sacrifice a life to save any number of people from headaches, it seems that one should be allowed to drive to the store to get some aspirin to relieve a single headache—even though driving to the store has some small probability of leading to a fatal car accident.

Lazar and Lee-Stronach (2019) are aware of these pitfalls, and they avoid them. They deny that the value of a headache or a cup of coffee is constant. Instead, they contend that the values produced by headaches and cups of coffee are bounded—the more of them they are, the less significant they are. If the total good that might be realized by a lesser good is bounded, we can explain why it would be wrong under conditions of certainty to trade off lesser goods for some superior good without positing any sort of infinite value and facing the difficulties that such infinite values engender. While we allow that their approach enjoys some important advantages over views that posit infinite values, we think that their proposal suffers from some significant problems. We'll present a series of objections and do our best to keep the model going in the face of these problems. We hope these objections will be instructive. The exercise is an important one: as our introductory remarks suggest, the bounded approach is one of the few options available for crafting a version of absolutism that avoids the threat of paralysis.[3]

The theory

Suppose that goods and bads, weighty or trivial, are tokened. The total value of an option will be determined by the value of these tokens. If we were to determine the total contribution of some type of good by, say, assigning the same value to each token and summing them up, this simple additive view would force us to countenance the possibility that a sufficient number of trivial goods (e.g., the pleasure of watching a few more seconds of a live football match) could conceivably contain more value than some significant good (e.g., the value of the life of a technician we could only save by unplugging equipment that must be powered on to continue the broadcast) unless we were to assign comparatively extreme values to these more significant goods (e.g., assigning each life infinitely more value than football match watching).[4] To avoid this dilemma, Lazar and Lee-Stronach tell us to think of the aggregate value of some goods as bounded so that the total value realized by tokens of this good might approach but never exceed some value.

Let's consider an example. It may be that one person's watching a football match for a few seconds is worth 1 utile, two people watching a football match for a few seconds is worth 1 ½ utiles (the marginal utility of the second person's watching being ½ of a utile), three people watching a football match for a few seconds is worth 1 ¾ utiles (the marginal utility of the third person's watching being ¼ of a utile), and the contributions of tokens of asymptotically decrease so that the maximum possible contribution of such tokens is bounded by 2 utiles. We can also have a similar structure for headaches, with the first person's headache having utility of negative 1, and the maximum (negative) possible contribution of such tokens being bounded by negative 2 utiles. We

[2] See Hausner (1952) and Hausner & Wendel (1952) for the classic mathematical work on lexically ordered utilities. See Hájek (2003) and Russell & Isaacs (forthcoming) for more recent philosophical overview.

[3] A salient alternative absolutist strategy is one that enjoins us to neglect certain small probability outcomes, allowing us to disregard the possibility of killing if it is small enough. For a sympathetic presentation of this style of decision theory see Monton (2019). For critical discussion see Hajek (2014) and Isaacs (2016).

[4] This parenthetical example is a variant of a famous case from Scanlon (1998).

can use these bounded structures to formulate a variety of absolutist views.[5] The key idea is that the value of this or that lesser good may never match or surpass the value of a token of some greater good.[6] The use of bounded value curves frees the absolutist from any commitment to the existence of goods of infinite value.

These bounded value curves are supposed to earn their keep by helping us formulate a plausible theory of moral decision making under risk.. Actions are then evaluated according to the sum of the values of their outcomes weighted according to their probability of obtaining if the action is performed. Suppose now that each token of some weighty good C is worth 5 and some trifling good c is such that its first unit is worth 1 and the total value of its tokens is bounded at 4. It clearly does not follow that it is always better to opt for a chance of a token of C than high probability of some quantity of tokens of c. To simplify, suppose that C and c are the only kinds of values in play and that S has two options, X and Y. X has a .1 chance of producing 1 token of C and 0 tokens of c and a .9 chance of producing 1 token of c and 0 tokens of C.[7] Y guarantees 0 tokens of C and a huge number of tokens of c with a value close to 4. The second option will be preferable.[8] The key idea is that even if the value of tokens of c is bounded below some quantity of tokens Q (perhaps 1) of C, there will always be a real number such that the result of multiplying the value of Q of C by that real number takes it below the upper bound for the value of cs.

In the discussion that follows, let's have one serious moral consideration C in play – saving lives. It will be helpful to have a few trifling goods and bads in play. Let's say moderately enjoyable but brief chocolate snacks, tasty and satisfying cheeseburger lunches, and mild headache avoidances are goods bounded below the value of 1 unit of C. Just to fix ideas let's suppose that 1 unit of C is worth 5, that 1 unit of each of the trifling goods is worth 0.1 and that the trifling goods are bounded at 4. Let us suppose further that mild headaches and have a negative contribution that is bounded at minus 4 and that 1 unit is worth minus 0.1. (We could obviously rescale things so that there were no negative numbers in play, but we follow Lazar and Lee-Stronach in using negative numbers in our presentation.)[9]

---

[5] Lazar and Lee-Stronach (2019, 98-99) define *extreme absolutism* as the view that there exists some kind of good such that a single token of this token is more valuable than any amount of any other kind of good, m*oderate absolutism* is the view that there are particular pairs of goods, C and c, such that any amount of the former is more valuable than any amount of the latter and *mild absolutism* is the view that there is some particular pairs of goods C and c such that some quantity of the form is more valuable than any amount of the latter. Our critical discussion will make trouble for all three versions. (For more on this taxonomy of absolutist theories, see Lazar (2018).)

[6] We realise that a few philosophers—e.g., Taurek (1997)— have said that killing many is no worse than killing one. Such philosophers might think that causing a lot of headaches is no worse than causing a few. But such ideas have no place in the current approach: the absolutist view under consideration seems to go in for a kind of aggregation according to which lots of headaches are worse than fewer (though of course never enough worse to make their avoidance worth a life).

[7] The ideology of 'in play' is one we borrow from Lazar and Lee-Stronach though we shall subject it to a little more critical scrutiny later.

[8] This structure can be replicated for each of extreme, moderate or mild absolutism (see fn. 5).

[9] Their account provides a nice model for mild absolutist intuitions (in the sense of fn 5). Here is one such setting: Two bullets, A and B are heading towards Sally, both would fatally wound her if the first to enter, A trails B. If A hits, Sally will die at t. If B hits, Sally will die at t plus a tenth of a second. Abraham is given the choice between pushing a red button, averting 10000 mild headaches with A hitting, and pushing a blue button, averting A hitting but not averting any headaches. For some short period of time, it does not seem that the delay will be worth giving up on the headache avoidance. When we save someone we delay their death? By how much? We have just seen that the answer to this will matter. If we don't delay their death by enough, then it won't be worth allowing the headaches after all. Delaying saving someone is shortening the amount one delays their death. But if we are absolutists at all, we should say that delaying their death by forty years is worth any amount of mild headaches. So we have a quantity (i.e., delaying their death) that stands in the 'weakly absolutely preferred' relation that Lazar and Lee-Stronach discuss. Certain quantities of it are not

Two notes of clarification are in order.

First: We have, naturally enough, conceived of chocolate snacks as trifling goods. But the situation arguably changes if chocolate snacks have buying power. Suppose snacks can be traded for yaks and yaks can save lives. Then it is far more tendentious to categorize snacks as trifling goods.[10] To control for this let us stipulate that the snacks, cheeseburger lunches, etc. are available in settings where they cannot be traded for other goods. Of course this is somewhat artificial, but it allows us to engage with the relevant issues by idealizing away an extra dimension of complexity.

Second: The idea that the value of some trifling good like tasty chocolate snacks is bounded has some superficial resemblance to familiar ideas about the diminishing marginal utility of tasty chocolate snacks to an individual. If we give a hungry individual a chocolate snack, that may benefit that individual quite a bit more than a second chocolate snack, which may in turn benefit them more more than a third chocolate snack. And if the individual is unable to trade the chocolate snacks for other goods or hand them over to other people, then the benefit of chocolate snacks will trail off quite quickly (and very quickly indeed if the shelf life is very short). The shape of that story has a superficial resemblance to the Lee-Stronach vision, but the latter seems to require quite different conceptual underpinnings. The reason why the benefit of chocolate snacks trails off for an individual is intuitively because the pleasure that they deliver trails off as the individual because less hungry and more engorged. But if the value of giving a chocolate snack to each of 10000 hungry individuals is rather less than double the value of giving a chocolate snack to each of 5000 individuals, that isn't going to be because the kind and quantity of pleasure that is delivered by the chocolate snack trails off. The justification of bounds will have to be different in that case. That is not to say that such bounds are out of the question. Indeed it is quite routine in economics to use decision theories with bounded utilities, which implies that the utility potential of chocolate snacks is bounded somehow or other. Utilizing bounded utilities is by no means eccentric. And so the key innovation here—that the value of trifling goods is bounded below certain other goods—deserves careful consideration.

Problem 1: Comparing Mild Goods and Bads

The theory is designed to vindicate standard absolutist comparisons between weighty and trifling considerations without making any chance of a weighty good automatically trump any trifling goods, and similarly without making any chance of a weighty bad automatically trump any trifling bads. But these features may come at the cost of making for bizarre results when it comes to choices between trifling goods or bads:

> A dodgy deity gives Abraham the following forced choice: Press button A and a million people on the West Coast of Country X will have a chocolate snack and one person on the East Coast of Country X will have a tasty cheeseburger meal or press button B and two million people in country Y will get a chocolate snack.

superior to certain quantities of headache avoidance but some quantities of it are superior to any quantity of headache avoidance.

[10] Certain of Lazar and Lee-Stronach's examples are arguably tendentious for related reasons. They write that "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." (p.40) If the well off people were able to use money to save lives, the calculation becomes considerably more complicated.

It is easy to imagine that the chocolate snacks get very close to their bound well before a million. But in that case the decision theory would seem to enjoin Abraham to prefer the million snacks with one cheeseburger to two million chocolate snacks. (Remember that we are assuming that trifling goods cannot be traded). This is bizarre if, as we might imagine, Abraham might have recognized that the first cheeseburger is not discernibly better or worse than the first tasty snack.

The challenge can be met. A natural fix is to say that the multiplication of one kind of trifling good has a knock-on effect on the value of another. And an obvious way to implement this thought is to have one grand category of trifling goods and a function from the number of tokens in this category to bounded values.

Problem 2: How Quickly Do the Value of Trifling Goods Trail Off?

We have said that the value of trifling goods is bounded. But how quickly does the value trail off? In our toy model, the value of a serious good was 5 and the value of chocolate snacks was bounded at 4. But that does not tell us how quickly the bound is approached as the quantity of chocolate snacks increases. We think there is an uncomfortable choice point here. Suppose one thought that it trailed off quite quickly, so that 50 chocolate snacks (distributed across 50 hungry individuals) already had a value of 3 and so 100 chocolate snacks (distributed across 100 hungry individuals) had considerably less than double the value. We can test such a theory by, inter alia, looking at its predictions about the comparative value of various risky choices whose possible outcomes all concern chocolate snacks. Here is one:

> Jones is given the choice between pushing two buttons. Button A has a one quarter chance of yielding tasty chocolate snacks for each of 100 individuals (and otherwise will do nothing). Button B has a one third chance of yielding tasty chocolate snacks for each of 50 individuals (and will otherwise do nothing).

It seems obvious that we should opt for button A. But if the value of chocolate snacks trails off at the kind of rate envisaged, we are enjoined to hit button B.

There is a different kind of vision, according to which the value of chocolate snacks trails off only as the numbers get *huge*.[11] Suppose for example, that the value of ten million hungry people each getting a chocolate snack is only just below double that of five million and only just below ten million times the value of one hungry person getting a tasty snack: the value of tasty snacks only trails off substantially as one approaches huge numbers. Whether results such as the above – with the example changed to appropriately huge numbers – are now more palatable is a good question. But in any case, this last approach has some distinctive costs vis-a-vis the kind of problem that Lazar and Lee-Stronach are trying to solve. It is paradigmatic of the kind of absolutism that they are trying to justify that ten million tasty snacks isn't worth killing someone for. But if ten million snacks is almost ten million times as good as one tasty snack, then it seems anything close to a one in ten million chance of running someone over is not a risk worth taking if the benefit from driving is getting a chocolate snack from the local grocery store. We thus risk the very kind of paralysis that

[11] One way to do this which is in keeping with the use of bounded utilities in economics though not really suggested by Lazar and Lee-Stronach's text is to discount the value of a hungry person enjoying a tasty snack only for episodes distant into the future. If the current hungry population of the cosmos is n, there is no discounting at all across possible situations in which the current number of tasty snacks is n, but for numbers of snacks greater than n, situations in which they are had will be discounted more at more as more and more temporally distant people enjoy them.

this theory is designed to avoid. Now of course the numbers in our toy model – where the trifling good is bounded at 4 and the serious good set at 5 are not set in stone. But it is not easy to see how to get satisfactory results. After all, increasing the gap between the serious good and the upper bound of the trifling good will only accentuate the problem just presented.

Problem 3: Sweeteners and Dampeners

Suppose saving a life is worth 5 and averting headaches is bounded at 4. Then no amount of mild headache avoidances should be preferred to saving a life. But adding a dampener to the saving of a life or adding a sweetener to the headache avoidance will change things.

> On Monday, the dodgy deity offers Abraham the choice between saving a life and averting a million mild headaches. As a good absolutist he chooses to save a life. On Tuesday he is offered a new choice: Save a life and cause ten nasty taste sensations or avoid a million mild headaches. A dampener – ten nasty taste sensations – has been introduced into the noble saving a life option. Abraham says: "I'd rather avert all those mild headaches and so I'll forgo the option of saving a life, now that it comes with ten nasty taste sensations." On Wednesday the dodgy deity offers Abraham another choice: Save a life or avoid a million headaches and get a million chocolate snacks handed out. "I'm throwing in the snacks to sweeten the 'averting a headache' option" says the deity. Abraham goes for the option with a sweetener.

Abraham doesn't seem like a very good absolutist. One would think that no absolutist worth their salt would be deterred from saving a life by the prospect of ten nasty taste sensations or by the extra enticement of chocolate snacks. But these conclusions are not easy to avoid in Lazar and Lee-Stronach's framework.

A move from the previous discussion might help with the Wednesday issue. Suppose the value of any trifling good was diluted by the presence of lots of them in such a way that tokens of trifling goods had a combined upper bound lower than saving one life. Then the averted headaches and snacks could not combine to beat the saving of a life. But this doesn't handle the Tuesday problem.

Here some further ingenuity will be needed to get an axiology that is friendly to absolutist instincts. Suppose the negative value of certain quantity Q of bad tastes, in combination with the value of saving a life, sometimes brings us under the threshold for headache avoidance. Then we will get the untoward result about dampeners. But we can in principle avoid the problem if combinations of trifling dampeners and a serious good can never take one below the upper bound for mild headache avoidance. There are a few ways to get this result. We can posit an invariant bound n for trifling goods and an invariant bound minus m for trifling bads and posit a value of saving a life that is greater than n + m. Then, by hypothesis, no amount of trifling dampeners, in combination with saving a life, will take one below the upper bound for trifling goods. In effect, this approach institutes a sufficiently big buffer zone between the upper bound of the trifling good and the value of a serious good in such a way that the (invariant contribution) of this or that number of trifling dampeners can never fill that buffer zone. Call this the Buffer Strategy. Alternatively, one can posit interaction effects: The value of this or that quantity of trifling goods or trifling bads is not constant

across worlds but is sensitive to what else is going on. This will involve giving up a picture according to which onee one gets the value of an outcome by adding the contributions from each moral consideration in favor of a more holistic picture where there is a function from token combinations to values that cannot be so decomposed. And with this more holistic approach at hand, we can imagine that the lower bound for saving a life in combination with any quantity of trifling bads was above four even while allowing trifling bads to have a more marked negative effect in other possible settings. Call this the Holistic Strategy.

Problem 4: Sweeteners and Dampeners in Situations of Uncertainty

> On Monday, the dodgy deity offers Abraham a choice: An 80 per cent chance of saving a life and 20 per cent chance of nothing of note vs. a 90 per cent chance of avoiding 10,000 mild headaches and 10 per cent chance of nothing of note. "I'll opt for the .80 chance of saving a life every time!" Abraham says. On Tuesday, the deity throws in a dampener into the first option: 80 per cent chance of saving a life and 20 per cent chance of 10 nasty taste sensations vs. 90 per cent chance of avoiding 10,000 mild headaches 10 per cent nothing of note. On Wednesday the deity instead throws a sweetener into the second option: 90 per cent chance of avoiding 10,000 mild headaches and 10 per cent chance of 10 chocolate snacks.

Is it really so absolutist friendly to think that the absolutist should be manipulated in and out of saving lives by sweeteners and dampeners in this way? The .80 shot at saving a life might be a matter of throwing a life jacket in the direction of someone who has capsized, where it is likely but not inevitable that it will save them. The sweetener might be a matter of having the opportunity of serving 10 chocolate snacks instead of throwing the life jacket. Are we really up for absolutists that go for the chocolate snacks here, at least when one is averting a lot of mild heachaches as well? Or the dampener might be the prospect that the person is not drowning and one will hit their head and cause a headache by throwing the life jacket.

Notice that the Holistic Strategy described at the end of the last section is irrelevant here, since the sweetener and dampers do not co-occur with saving lives or averting headaches in any of the possible worlds of interest. How about the Buffer Strategy? Clearly that will not stop Trifling Sweeteners and dampeners making a constitutive difference. Suppose the value of saving a life is 10, the value of tasty snacks is bounded at 4 and the disvalue of mild headaches is bounded at 10. It will still be easy to construct pairs of cases where the chance of a sweetener or dampener makes a difference to whether one takes on a risk of killing: we can use the above vignette with percentages tweaked so that sweetening by a chance of snacks or dampening with a chance of mild heachaches makes a difference. This is a bullet that the current brand of absolutism will just have to bite. We leave it to readers to ponder whether this sacrifices too much of the absolutist spirit that the theory is intended to preserve.

Problem 5: Multiple Bounds

In the simple models that Lazar and Lee-Stronach present as a toy starting point, there are weighty considerations that are unbounded and trifling ones that are bounded. But the situation may be

more complex. Even leaving aside those considerations from decision theory that push towards overall bounded utility functions, the absolutists likely need to have multiple layers of unacceptable trade-offs. This will incline one to a more complex theory. Suppose no amount of mild headache avoidance will compensate for an amputation, and no amount of amputations will compensate for killing someone. (If amputation does not seem to the reader like a suitable candidate intermediate bad of the relevant kind, then we invite them to choose their own favorite toy example: A loss of a friendship, the loss of an ability to appreciate art, the killing of a squirrel etc. etc.) One might think that one can accommodate these absolutist intuitions by headache avoidance by instituting multiple bounds. Perhaps, for example, the disvalue of headaches is bounded at 4, the disvalue of amputations is bounded at 1000 and a single death has a disvalue of 1500. But we cannot really accommodate absolutist intuitions by this strategy. Consider:

> On Monday a dodgy deity gives Abraham a forced choice: Push button A and avert 1,000 amputations or push button B and avert 10 mild headaches. Abraham avoids the amputations. On Tuesday a dodgy deity gives Abraham a new forced choice with the same structure: Push button A and avert 1,000 amputations or push button B and avert 10 mild headaches.

With the two boundaries in place Abraham may well be obliged to prevent 10 mild headaches rather than avert 1000 amputations – a ridiculous result. But if the value of avoiding amputations is bounded, it is unavoidable that the additional value of avoiding 1000 extra amputations may be minuscule. If the difference between 0 headaches and 1 mild headache is constant across possibilities, it will be greater than the difference between n and n + 1000 amputations for some value of n. This hardly vindicates the idea that no number of mild headaches can compensate for an amputation – it's a situation where no number of amputations can compensate for a headache! Thus, in the presence of a more complex hierarchy of goods, it is much more challenging to work out an acceptable version of Lazar and Lee-Stronach's framework.

One defensive strategy here is to posit a kind of holistic strategy where suitable interaction effects control for the difficulty. As a toy model, imagine that if there have been n amputations, then the most disvalue that mild headaches can produce is an amount less than the difference between n and n + 1 amputations. In effect, the more amputations there have been the less disvalue mild headaches contribute. Assuming this interaction effect, the ridiculous result described above is avoided. But there is a revenge problem. Consider the following:

> A dodgy deity gives Abraham the following choice. Push button A and there is a .8 chance of 1001 amputations and a .8 chance of nothing. Push button B and there is a .8 chance of 1000 amputations and a .2 chance of ten mild headaches (and no amputations).

Assuming 1000 amputations is sufficiently close to the relevant upper bound (and if not, pick a number that is), the strategy just adumbrated recommends pushing button B in this scenario: when the ten mild heachaches occur they are not bundled together in a world with myriad amputations and so there is not the signature interaction effect posted by the holistic strategy just described. But we submit this is, once again, not a result that will be palatable to absolutists. The kind of absolutist who is gripped by the thought that no amount of mild headaches is as bad as an amputation is not likely to be the kind of absolutist that is comfortable with pressing button B in the scenario above.

Problem 6: What are outcomes?

Standard decision theory distributes values and probabilities over outcomes. It calculates the value of an act in terms of the sum of the products of the values of each outcome and the probabilities of those outcomes conditional on the act being performed.[12] But what is an outcome? In some places, Lazar and Lee-Stronach talk as if, for the purpose of moral decision theory, we can treat outcomes as states of a world that encodes at least everything of moral relevance about a world.[13] This would be very much in keeping with the standard way of thinking about outcomes as the kinds of states to which maximally fine-grained utility assignments can be made. But we are not sure they have thought through the consequences of this standard way of thinking about outcomes, and indeed have a tendency in places to think about outcomes in a different way.

Suppose to simplify, that there are just two goods, tasty and satisfying cheeseburger meals and loving friendships (one trifling and one weighty), and no bads, and the contribution of cheeseburgers and friendships is a matter of their quantity. One might then naturally think that the value of a possible world is a function of the number of cheeseburger meals and loving friendships that occur in this world. What this will mean is that knowledge about the past will have a significant bearing on one's estimation of the value of one's actions:

> A dodgy deity gives Abraham a choice: .01 percent chance of a loving friendship or creating 10000 tasty cheeseburger meals. Abraham is inclined to go for the cheeseburgers. But then the deity reveals some news: In the past there have been twenty million cheeseburger meals. "Oh," Abraham says, "I thought most likely there hadn't been any tasty cheeseburger meals yet. Now I'll take my chances going for the friendship." The revelation changes what Abraham prefers. The utility contribution of extant cheeseburger meals is already close to the upper bound for that consideration. It's not worth foregoing even the low chance of a friendship for any number of cheeseburger meals.

This sensitivity to the past is inevitable if the value of an outcome is anything like the value of the possible world that is actualized by that outcome. And it seems very strange to go in for an axiology where the past matters in *this* kind of way.[14] In any case, an axiology with this kind of past-sensitivity is not faithful to Lazar and Lee-Stronach's guiding vision. It bears emphasis, in this connection, that they talk about outcomes in a very different way. One of the main challenges they anticipate to their theory has to do with potentially contrasting computations for a complex series of actions when the evaluation is done for the complex as opposed to the individual actions that make it up. We shall be directly addressing what they have to say about this "Act/Campaign Challenge" in a later section. But what matters in the present context is that their discussion evinces unclarity about the operative notion of outcome. In one place, for example, they imagine that X is the value of a "single token on its own" of a certain type of moral consideration c and that there is a sequence of acts in which "a single token of c is at stake" (Lazar and Lee-Stronach, 2019, p. 106). They then say that the

[12] Alternative decision theories do exist, though engagement with such alternatives would take us too far afield. For an overview of some alternatives (particularly risk-weighted expected utility theory) see Buchak (2013).

[13] For example, they refer to a "decision problem in which φ and ψ are actions, and A and B are possible states of the world." (Lazar and Lee-Stronach, 2019, p. 101).

[14] For a closely related argument in population ethics, see Wilkinson (2022).

"aggregate moral utility of those ten instances of c will be 10X" (ibid). They then say that by contrast "if ten tokens of c are at stake in a single act, ψ, they must together weigh less than 10X" (ibid). (Thinking about things this way the act/campaign challenge then emerges very sharply: while a single act of creating multiple cheeseburgers cannot generate the value of a loving friendship, a sequence of such acts can do so.) However, it is not easy to make sense of what is going on here. Suppose one performs the tenth instance of creating an X token. Then one is in a world where nine tokens have already been produced. But presumably the axiology will tell us that, *ceteris paribus*, the difference between a world with 9 units and that with 10 units is less than X, and if there are other units produced elsewhere that will also be relevant. In the context of those passages the authors seem to have a very different notion of outcome in play than 'totally morally relevant state of the world'.

To try to accommodate these remarks, we can imagine an axiology that worked like this: To find the value of an action, just look at the causal consequences of the action, and treat the total value of the goods and bads instantiated by a particular future as insensitive to the goods and bads instantiated in the past. Of course, as usual, the causal consequences of the action will depend on the state of the world one is in: if one gives money to someone already rich that will have different causal consequences to giving money to someone who is poor. Similarly, whether a cheeseburger is tasty and satisfying or not will depend *inter alia* on what else has been consumed recently. (Thus the approach being envisioned here need not ignore the mundane kind of diminishing marginal utility noted in our initial remarks.) And for certain goods and bads, whether they are instantiated at all will *constitutively* depend on the past: if one goes to the movies, that may count as keeping or breaking a promise depending on past promises. The key thought is that it is the relative quantities of goods and bads in the future that are determinative of the value of the action, with the quantity of various goods and bads in the past being irrelevant. Note that this is compatible with the kind of interactive holism presented earlier, except that interaction effects relevant to the value of an action are confined to the goods and bads in the future of an action. For any actual action, its value will supervene on the distribution of goods and bads in its future (and for any counterfactual action, its value will supervene on its counterfactual future).

On this way of thinking it may not matter to the value of a present action whether there have been a million cheeseburgers meals (at least insofar as this past distribution has no causal bearing on whether future cheeseburger meals are tasty and satisfying), since past meals are in the past and only future goods and bads matter to the value of that action.

However, even this future-theoretic conception does not quite seem to be what the authors want. If one chooses to token a c, that may not be causally irrelevant to actions that produce a token of c down the line. But they want a decision theory where only "a single token is at stake" and "in play" when performing each act. But these notions of at stake and in play are not ones that we find in standard decision theory. When one considers various actions there are various futures that are possible each of which will involve yourself and other agents doing things. Each of those futures will no doubt contain multiple tokens of c. What exactly are we to mean by "only one token is in play" or "only one token is at stake". The ideology of outcomes in their theory remains very obscure and would need to be precisified before any careful discussion of the value of sequences of actions can proceed.

Here is one idea: Let the constrained zone of influence of an action be the causal consequences fanned out until they hit other decisions to act. Upshots of those other acts are beyond the constrained zone of influence. One might then run the axiology by having it evaluate actions by looking at combinations of goods and bads in the constrained zones of influence.The

value of an act is given by adding the goods and bad within the constrained zone of influence.[15] This would give them the result they want in the ten act case. When S decides to produce a token of X there is a second decision standing between S and a second token. So the second token is beyond the constrained zone of influence relevant to the first decision. The sum total of tokens of X within its constrained zone of influence is one. It bears emphasis how alien this way of thinking about things is. Suppose one is deciding whether to produce a token of X. One expects one's decision to have some influence on whether a second decision about a token of X is positive: one thinks it more likely the second decision is positive conditional on the first being positive. The kinds of utility functions standardly deployed in decision theory tell us to take such facts of expected influence into account. But the current proposal treats such lines of influence as irrelevant.

Problem 7 Telling People what to do

We can dramatize some themes from the last section

> Abraham has a choice. He can tell Frank to save a life and is .99 that he will be successful (Frank generally does what he is told) and .01 he will do nothing. Alternatively he can issue a group instruction to a very large group of two million people to each see to it that one headache gets averted. He is certain that at least one million of them will follow the instruction and that the instruction will, if not followed, have no impact of note at all. The scale for headache aversion is as before.

What is Abraham to do? If we take the "ten acts, 10 times X" heuristic seriously, it looks like one is guaranteed to produce vastly more utility by instructing the members of the big group to each see to it that a headache is averted. (Moreover some of the complexities of deciding whether to count complex acts by oneself as a single big act or a series of small acts are neither here nor there.) So if Abraham looks ahead to the resultant utilities, he will issue the instruction: "Avert mild headaches!" That is hardly friendly to absolutist sensibilities. Alternatively, if we take the "restricted zone of tolerance approach" the bad result is avoided at the price of indifference: "There is nothing in the restricted zone of tolerance either way so it is of no ethical importance to Abraham which instruction to go for". Neither option seems appetizing to the absolutist.

Problem 8: The agglomeration of risk

[15] There are also intermediate views of course: For the purpose of evaluating an act, one could weight goods and bads within the constrained zone of influence *far more heavily* than aspects of the future that lie outside the constrained zone of influence without going all the way to a theory that is utterly preoccupied with that zone for the purposes of act evaluation. (This could then interact with absolutism in interesting ways. For example, a saving of a life within the constrained zone of influence may have a value above the upper bound of trifling goods, but a saving of a life outside the constrained zone may be discounted to a point below that. So, for example, taking a .9 risk of inciting a potential axe murderer to kill may be compensated for by trifling goods in a way that killing someone yourself cannot, and that the .9 risk still count as very weighty.) This kind of view seems more palatable than the restricted zone of influence view discussed in the next, though less faithful to Lazar and Lee Stronach's talk about how many tokens are at stake or in play.

The view seems to tell us some very strange things about the agglomeration of risk. We might be offered a series of gambles where the potential prize is the averting of headaches and the possible loss is the loss of a life. Moreover, the risk to life falls below the threshold for each gamble and yet it's certain that if we take all of the gambles a life will be lost. Imagine an absolutist were offered this gamble:

> A dodgy deity has a prisoner in a cage. To play, you have to press a button that either frees or kills the prisoner. There is a one sixth chance that you will kill the prisoner by pushing the button (since the button is hooked up to a six-shooter pistol with all but one chamber empty. You know that you'll never know whether a prisoner has been freed or killed by a particular button pushing (the pistol has a silencer and the cage is soundproofed) . You know that each time you press a button, it's a certainty that an astronomical number of headaches will be averted. You have six opportunities to play the game, with each opportunity separated by a 30 second interval, beginning at noon.[16]

Under the current setup, since the probability of death if you press at noon is ⅙, you should press the first button. The value of the headaches, though bounded, exceeds ⅙ of the value of a life. For the same reasons you should press the button at 30 seconds after noon, the button at a minute after noon and so on. At each time, the chance that hitting the button will kill is the same. (For example, if it is hit a sixth time, it is hit at a point where it is 5/6th likely that the person is already dead.) But it is crucially important that your lethal activity is spread over six acts. Suppose for example that at each time you had the alternative option of hitting a button hooked up to a fully loaded pistol that would avert twelve times the number of headaches of the other button. (We can institute a three minute delay so that this option will certainly not hasten death vis a vis a sequence of six button pushings.) Hitting *that* button would be utterly impermissible in the current framework (even though it generates double the headache aversion benefit of the prolonged activity). But it seems utterly wrongheaded that so much should depend on button-pushing logistics. If it's wrong to press the lethal button, then it's wrong to play six rounds of Russian Roulette in sequence. (Similarly, if it's wrong to press the Russian Roulette button six times in a row, then it's wrong for six different people to coordinate to each push it once.) It's simply not reasonable to think that so much of moral value depends on how many acts are performed to effect some consequence – that sensibility is utterly alien to normal moral theorizing.

The challenge here is intimately connected to what Lazar and Lee-Stronach call the "act-campaign" challenge. The general issue, as they see it, is when it is decision-theoretically acceptable to assess acts in a sequence individually rather than to bundle a sequence together and consider it as one single, temporally extended act. Their proposal on this matter is this: "when a risky act is causally sufficient to realize some expected good it can be considered in isolation from the campaign of which it is part, and assessed as permissible or impermissible. When one risky act

[16] It is useful to distinguish this example from a diachronic example in which there are six buttons hit one at a time, six prisoners, and exactly one of the buttons will kill but one does not know which. (As with the example in the main text, one does not know the result of one's pushing a particular button.) Aboodi, Borer and Enoch (2008) sketch an absolutist treatment that defuses this last case. Their idea is that deontological constraints relate actors to particular individuals. So if the key threshold of risk is t, then in the case at hand, the key question is whether the risk to prisoner A is above the risk threshold, whether the risk to B is above the threshold. As there is no deontological obligation to the fusion of the six prisoners, the (maximal) risk of killing one of them is neither here. Aboodi et al's tweak on absolutism is neither here nor there when it comes to our example, since there is only a single candidate victim in play.

depends on others to realize its expected good, then they must be assessed together." (Lazar and Lee-Stronach, 2019, p. 108). On this proposal we get the result that each pushing of the button is permissible because each alleviates a large number of headaches at only 1/6th risk of causing death. Thus, on their vision, the whole sequence is permissible. But this is exactly the result that we have labelled as wrongheaded in this kind of case.

It bears emphasis that the challenge here is not fundamentally a challenge to Weak Agglomeration, the principle that if you ought to do X at t_1 and you ought to do Y at t_2 you ought to do X at t_1 and Y at t_2. The strongest prima facie challenge to that principle is arguably raised by bundling together *synchronic* actions (where $t_1=t_2$), though even in that setting the challenge is resistable, even by one of an absolutist bent.[17] In the setting of this diachronic sequence, it is easy to avoid Weak Agglomeration: we can say that one ought to perform a sequence iff one ought to perform each act in that sequence as judged by one's favorite theory. And as we have seen, Lazar and Lee-Stronach's favorite theory seems to prescribe the sequence of hitting the button six times. The problem isn't that this violates Weak Agglomeration. The problem is that it accords unacceptable moral import to the distinction between hitting the Russian Roulette button six times and hitting a lethal button once.

Conclusion

Our overall assessment should be clear enough. While it certainly has some initial appeal, the strategy of using bounded utilities to save absolutism from practical paralysis faces a range of challenging obstacles. For what it's worth, we are sceptical that they can be overcome in a way that vindicates absolutist instincts.

References

Aboodi, R., Borer, A. and Enoch, D. 2008. "Deontology, Individualism and Uncertainty: A Reply to Jackson and Smith," *The Journal of Philosophy* 105: 259-272.

Black, D. 2020. "Absolute Prohibitions Under Risk," *Philosophers' Imprint* 20 (20): 1-26.

Buchak, L., 2013. *Risk and Rationality*. Oxford University Press.

Dorsey, D. 2009. "Headaches, Lives and Value," *Utilitas* 21: 36-36.

[17] Recall the two skier case in Jackson and Smith, 2006, p. 276: One can shoot either or both of two skiers A and B. In each case one is likely killing a suitably malevolent skier but there is risk that if one shoots, one will shoot an innocent person. For a given threshold of acceptable risk t, one can set things up so that the chance that A is innocent is below t, the chance that B is innocent is below t, but the chance that one of A and B are innocent is above t. On a naive way of running absolutism, you may end up claiming that one ought to kill A, one ought to kill B, but it's not the case that one ought to kill both. As Black (2020) points out though, there is plenty of room for a threshold inspired view to avoid violating Weak Agglomeration: If one evaluates an act by whether it surpasses the risk of doing a serious wrong, then the act of killing both skiers will be ruled out. And if killing just A and killing just B are tied, then both are arguably permissible, and so neither is something one ought to do. But avoiding Weak Agglomeration failure does not thereby avoid getting strange results for sequences of acts.

Hausner, M., 1952. *Multidimensional Utilities*. Rand Corp.

Hausner, M., Wendel, J.G., 1952. "Ordered vector spaces". *Proc. Amer. Math. Soc.*3 (6), 977–982. http://dx.doi.org/10.2307/2031746.

Hájek, Alan. 2003. "Waging War on Pascal's Wager." *Philosophical Review* 112 (1). Duke University Press: 27–56.

Hájek, Alan (2014). Unexpected Expectations. *Mind* 123 (490):533-567.

Isaacs, Yoaav (2014). "Duty and knowledge". *Philosophical Perspectives* 28 (1):95-110.

Isaacs, Yoaav (2016). "Probabilities cannot be rationally neglected". *Mind* 125 (499)759-762.

Jackson, Frank & Smith, Michael (2006). "Absolutist Moral Theories and Uncertainty". *Journal of Philosophy* 103 (6):267-283.

Kagan, S. 1988. "The Additive Fallacy," *Ethics* 99/1: 5–31.

Lazar, Seth (2018). "Limited Aggregation and Risk". *Philosophy and Public Affairs* 46 (2):117-159.

Lazar, Seth & Lee‐Stronach, Chad (2019). "Axiological Absolutism and Risk". *Noûs* 53 (1):97-113.

Monton, Bradley (2019). "How to Avoid Maximizing Expected Utility." *Philosophers' Imprint* 19: 1-25.

Norcross, Alastair (1997). "Comparing Harms: Headaches and Human Lives". *Philosophy and Public Affairs* 26 (2):135-167.

Russell, Jeffrey Sanford & Isaacs, Yoaav (forthcoming). "Infinite Prospects". *Philosophy and Phenomenological Research*.

Scanlon, Thomas (1998). *What We Owe to Each Other*. Belknap Press of Harvard University Press.

Taurek, John (1977). "Should the numbers count?" *Philosophy and Public Affairs* 6 (4):293-316.

Wilkinson, Hayden (2022). "In Defense of Fanaticism". *Ethics* 132 (2): 445-477.