In defence of fanaticism

Hayden Wilkinson

Global Priorities Institute | September 2020 (updated January 2021)

GPI Working Paper No. 4-2020

In defence of fanaticism

Hayden Wilkinson*

January, 2021

Comments welcome: hayden.wilkinson@anu.edu.au

Abstract: Consider a decision between: 1) a certainty of a moderately good outcome, such as one additional life saved; 2) a lottery which probably gives a worse outcome, but has a *tiny* probability of some vastly better outcome (perhaps trillions of additional blissful lives created). Which is morally better? By expected value theory (with a plausible axiology), no matter how tiny its probability of the better outcome, (2) will be better than (1) as long that better outcome is good enough. But this seems *fanatical*. So you may be tempted to abandon expected value theory.

But not so fast - denying all such fanatical verdicts brings serious problems. For one, you must reject either that moral betterness is transitive or even a weak principle of tradeoffs. For two, you must accept that judgements are either: inconsistent over structurally-identical pairs of lotteries; or absurdly sensitive to small probability differences. For three, you must accept that the practical judgements of agents like us are sensitive to our beliefs about far-off events that are unaffected by our actions. And, for four, you may also be forced to accept judgements which you know you would reject if you simply learned more. Better to accept fanaticism than these implications.

*School of Philosophy, Australian National University

1 Introduction

Suppose you face the following moral decision.

> **Dyson's Wager**
>
> You have $2,000 to use for charitable purposes. You can donate it to either of two charities.
>
> The first charity distributes bednets in low-income countries in which malaria is endemic.[1] With an additional $2,000 in their budget this year, they would prevent one additional death from malaria. You are certain of this.
>
> The second charity does speculative research into how to do computations using 'positronium' - a form of matter which will be ubiquitous in the far future of our universe. If our universe has the right structure (which it probably does not), then in the distant future we may be able to use positronium to instantiate all of the operations of human minds living blissful lives, and thereby allow morally valuable life to survive indefinitely long into the future.[2][3] From your perspective as a good epistemic agent, there is some tiny, non-zero probability that, with (and only with) your donation, this research would discover a method for stable positronium computation and would be used to bring *infinitely many* blissful lives into existence.[4]

[1] I have in mind the Against Malaria Foundation. As of 2019, the charity evaluator GiveWell estimated that the Against Malaria Foundation prevents the death of an additional child under the age of 5 for, on average, every US$3,710 donated (GiveWell 2020). Including other health benefits, a total benefit *roughly* equivalent to that is produced for, on average, every US$1,690 donated. Of course, in reality, a donor can never be *certain* that their donation will result in an additional life saved. This assumption of certainty is for the sake of simplicity.

[2] Dyson (1981) was the first to suggest positronium as a medium for computation and information storage. This follows Dyson (1979), wherein it is argued that an infinite duration of computation could be performed with finite energy if the computation hibernates intermittently, and if the universe has a particular structure. Tipler (1986) suggests an alternative method which may work if the universe has a different structure. Sandberg (n.d.) argues that both Dyson and Tipler's proposals are unlikely to work, as our universe appears to match neither structure. Nonetheless, it is still epistemically *possible* that the universe has the right structure for Dyson's proposal. And possibility is sufficient for my purposes.

[3] Would such artificially-instantiated lives hold the same moral value as lives led by flesh-and-blood humans? I assume that they would, if properly implemented. See Chalmers (2010) for arguments supporting this view. And note that, for the purposes of the example, all that's really needed is that it is epistemically *possible* that the lives of such simulations hold similar moral value.

[4] I have deliberately chosen a case involving many separate lives rather than a single person's life containing infinite value. Why? You might think that one individual's life can contribute only some bounded amount of

What ought you do, morally speaking? Which is the better option: saving a life with certainty, or generating a tiny probability of bringing about infinitely many future lives?

A common view in normative decision theory and the ethics of risk - *expected value theory* - says that it's better to donate to the speculative research. Why? Each option has some probability of bringing about each of several outcomes, and each of those outcomes has some value, specified by our moral theory. Expected value theory says that one option is better than another if and only if it has the greater probability-weighted sum of value - the greater *expected value*.[5] Here, the option with the greater expected value is donating to the speculative research (at least on certain theories of value - more on those in a moment). So perhaps that is what you should do.

That verdict of expected value theory is counterintuitive to many. All the more counterintuitive is that it can still be better to donate to speculative research *no matter how low* the probability is (short of being 0)[6], since there are so many blissful lives at stake. For instance, the odds of your donation actually making the research succeed could be 1 in 10^{100}. (10^{100} is greater than the number of atoms in the observable universe). The chance that the research yields nothing at all would be 99.99... percent, with another 96 nines after that. And yet expected value theory says that it is better to take the bet, despite it being *almost guaranteed* that it will actually turn out worse than the alternative; despite the fact that you will almost certainly have let a person die for no actual benefit. Surely not, says my own intuition. On top of that, suppose that $2,000 spent on preventing malaria would save more than one life. Suppose it would save a billion lives, or any enormous finite value. Expected value theory would say that it's still better to fund the speculative research - expected value theory says that it would

value to the value of the world as a whole - you might prefer for 100 people to each obtain some finite value than for one person to obtain infinite value. But whether this verdict is correct is orthogonal to the issue at hand, so I'll focus on large amounts of value spread over many people.

[5]Note that expected *value* is distinct from the frequently-used notion of expected *utility*, and expected value theory distinct from expected *utility* theory. Under expected utility theory, *utility* is given by some (indeed, *any*) increasing function of value - perhaps a concave function, such that additional value contributes less and less additional utility. The utility of an outcome may even be bounded, such that arbitrarily large amounts of additional value contribute arbitrarily little additional utility. Where expected value theory says that a lottery is better the higher its expected *value*, expected utility theory says that it is better the higher its expected *utility*. And, if the utility function is bounded, then the expected utilities of lotteries will be bounded as well. As a result, expected utility theory can avoid the fanatical verdict described here. But, if it does, it faces the objections raised in Sections 4, 5, and 6. Where relevant, I will indicate in notes how the argument applies to expected utility theory.

[6]I'll assume throughout that probability takes on only real values from 0 to 1.

be better to sacrifice those billion or more lives for a minuscule chance at the infinitely many blissful lives (and likewise if the number of blissful lives were finite but still sufficiently many). But endorsing that verdict, regardless of how low the probability of success and how high the cost, seems *fanatical*. Likewise, even without infinite value at stake, it would also seem fanatical to judge a lottery with sufficiently tiny probability of arbitrarily high *finite* value as better than getting some modest value with certainty.

Fanatical verdicts depend on more than just our theory of instrumental rationality, expected value theory. They also depend on our theory of (moral) value, or *axiology*. Various plausible axiologies, in conjunction with expected value theory, deliver that fanatical conclusion. Foremost among them is *totalism*: that the ranking of outcomes is determined by the total aggregate of value of each outcome; and that this total value increases linearly, without bound, with the sum of value in all lives that ever exist. By totalism, the outcome containing infinitely many blissful lives is indeed a *much* better one than that in which one life is saved. And, as we increase the number of blissful lives, we can increase *how much better* it is without bound. No matter how low the probability of those many blissful lives, the expected total value of the speculative research is greater than that of malaria prevention. (Likewise, even if there are only finitely many blissful lives at stake, for any tiny probability there can be sufficiently many of them to make the risky gamble better than saving a life with certainty.) But this problem isn't unique to totalism. When combined with expected value theory, analogous problems face most competing axiologies, including: averageism, critical-level views, prioritarianism, pure egalitarianism, maximin, maximax, and narrow person-affecting views. Those axiologies each allow possible outcomes to be unboundedly valuable, so it's easy enough to construct cases like Dyson's Wager for each.[7] And some - namely, critical-level views and prioritarianism - already deliver the same result as totalism in the original Dyson's Wager. In this paper, I'll focus on totalism, both to streamline the discussion and because it seems to me far more plausible than the others.[8] But suffice it to say that just about any plausible axiology can deliver fanatical verdicts when combined with expected value theory.

In general, we will sometimes be led to verdicts that seem fanatical if we endorse *Fanaticism*.[9]

[7]For instance, take (standard, welfarist) averageism. A population containing at least one blissful life of infinite (or arbitrarily long) duration will have average value greater than any finite value we choose. And so, to generate an averageist analogue of Dyson's Wager, we can substitute an outcome containing this population for the outcome of arbitrarily many lives in the original wager.

[8]Each of the other axiologies listed falls prey to devastating objections. See Arrhenius (2000), Huemer (2008), Greaves (2017), and chapters 17-19 of Parfit (1984).

[9]This use of the term 'fanaticism' seems to originate with Bostrom (2011) and Beckstead (2013: chapter 6)

Inversely, to succeed in avoiding fanatical verdicts, our theory of instrumental rationality and our axiology must not imply Fanaticism.

> *Fanaticism*: For any tiny (finite) probability $\epsilon > 0$, and for any finite value v, there is some finite V that is large enough that L_{risky} is better than L_{safe} (no matter which scale those cardinal values are represented on).
>
> L_{risky}: value V with probability ϵ; value 0 otherwise
>
> L_{safe}: value v with probability 1

The comparison of lotteries L_{risky} and L_{safe} resembles Dyson's Wager: one option gives a slim chance of a potentially astronomical value V; the other a certainty of some modest value v. But, here, V need not be infinite, in case you think infinite value impossible. And, with some minor assumptions (see Section 3), Fanaticism in this form implies the fanatical verdict in Dyson's Wager. Likewise, to reject the fanatical verdict in Dyson's Wager, we must reject Fanaticism.

Note that Fanaticism is quite a strong claim. As defined here, it requires that the ranking of L_{risky} above L_{safe} holds not only when the number of lives in each outcome are proportional to V, v, and 0. It must hold whenever outcomes can be *cardinally represented* with those values. Recall that cardinal representations of value[10] are unique only up to positive affine transformations - two outcomes represented by 0 and v on one scale could instead be represented by $0 \times a + b$ and $v \times a + b$ (for any positive a and real b). Conversely, an outcome that contains many happy lives might still be represented cardinally with value 0. So Fanaticism doesn't apply only to risky lotteries in which some possible outcome contains zero valuable lives, or zero value on net. It also applies to lotteries that can be represented as L_{risky} and L_{safe} even though every one of their outcomes contains enormous numbers of blissful lives, or enormous amounts of suffering, as long the differences in value between those outcomes are proportional to 0, v, and V.

Given how strong and how counterintuitive Fanaticism is, you might think it easy to reject. And many philosophers and other thinkers have done so, rejecting either Fanaticism or similar principles. For instance, Bostrom (2009) presents a compelling *reductio ad absurdum* for fanatical

(however Beckstead uses the term 'Fanaticism' for a similar claim specific to infinite values and instead uses 'Recklessness' for a claim more akin to my version of Fanaticism.). My formulation is slightly stronger than each of theirs but also, unlike theirs, applicable even if infinite total value cannot exist. For discussion of whether outcomes with infinite moral value are possible and how we might coherently compare them, see Bostrom 2011; Askell 2018; Wilkinson 2020; Wilkinson n.d..

[10] Equivalently, these are representations of value on *interval scales*.

verdicts in the prudential context. Bostrom (2011), Beckstead (2013), and Askell (2019) treat even a weak form of (moral) Fanaticism as a *reductio* for moral theories. Others propose theories of rationality with the expressed purpose of avoiding fanatical verdicts. For instance, some propose that we simply ignore outcomes with small enough probabilities (e.g., D'Alembert 1761; Buffon 1777; Smith 2014[11]; Monton 2019). Others insist that we maximise not expected moral value but instead the expected *utility* of outcomes (given by some increasing function of an outcome's value), and that the correct utility function is bounded above so as to keep the expectation of utility bounded as well (e.g., Arrow 1971: 64).

Meanwhile, there are few defenders of Fanaticism or, more broadly, of fanatical verdicts in cases similar to Dyson's Wager. Notable examples include Pascal (1669) himself, Parfit (1984: §3.27), and Hájek (2014). And even they most often endorse such verdicts only because they are a consequence of expected value theory, not because they see good independent justification for them. I suspect that even many diehard adherents of expected value theory are uncomfortable with the fanatical verdicts supplied by their theory.

This situation is unfortunate. There are compelling arguments in favour of Fanaticism that do not rely on expected value theory, and so we have good reason to accept it even if we reject that particular theory. If we do not, we face disturbing implications.

The paper proceeds as follows. Section 2 addresses some common motivations for rejecting Fanaticism. Section 3 introduces the necessary formal framework for what follows. Sections 4 through 6 present arguments in favour of Fanaticism, each premised on weaker claims than expected value theory, and each (by my reckoning) more compelling than the last. The first is a basic continuum argument. The second is that, to deny Fanaticism, we must accept either what I'll call 'scale dependence' or an absurd sensitivity to arbitrarily small differences in probability. And the final nails in the coffin are an updated version of Parfit's classic Egyptology Objection and what I'll call the *Indology Objection*, by which those who deny Fanaticism must make judgements which appear deeply irrational. Section 7 is the conclusion.

[11]Smith's proposal can be interpreted in two different ways, only one of which rules out Fanaticism. By the other interpretation, which Smith prefers, we still ignore events with probability below some threshold but, in any lotteries over finitely many different outcomes, that threshold is set below the probability of the least probable outcome. This is compatible with Fanticism while still avoiding the problems with which Smith is more concerned: counterintuitive verdicts in the St Petersburg and Pasadena games.

2 Intuitions against fanaticism

There are various arguments (that can be directed) against Fanaticism that appear in the literature (see D'Alembert 1761; Buffon 1777; Smith 2014; Schwitzgebel 2017; Monton 2019). And there are various counterarguments as well (see Hájek 2014; Isaacs 2016; Monton 2019; and Hájek n.d. for some of the most compelling).

But there is one extremely simple argument against it that, I suspect, is the most persuasive of all: Fanaticism is just frightfully counterintuitive that it must be false.[12] It seems so very implausible that I should give up a certainty of a good payoff for a *tiny* chance of something better, no matter how tiny the chance. Rather, intuition says that an event with such a tiny chance can never tip the scales against the sure payoff.

But the psychology literature tells us that our initial intuitions about matters of probability are often misguided. We frequently commit the Conjunction Fallacy (Tversky & Kahneman 1983), the Gambler's Fallacy (Chen *et al.* 2016), the Hot Hand Fallacy (Gilovich *et al.* 1985), and the Base Rate Fallacy (Kahneman & Tversky 1982). Perhaps worst of all, we are prone to simply ignoring facts of probability - many of us choose based entirely on how we anticipate things will *actually* turn out, with no regard for risks of loss or chances of gain (Baron *et al.* 1993; Gurmankin & Baron 2005). Even when we do account for probability, we often radically overestimate some probabilities due to availability bias (Tversky & Kahneman 1974), and underestimate others out of indefensible optimism (Hanoch *et al.* 2019).

And our intuitions are no better with decision-making in the face of low-probability events. For instance, jurors are just as likely to convict a defendant based primarily on fingerprint evidence if that evidence has probability 1 in 100 of being a false positive as if it were 1 in 1,000, or even 1 in 1 million (Garrett et al. 2018). In another context, when presented with a medical operation which posed a 1% chance of permanent harm, many respondents considered it no worse than an operation with no risk at all (Gurmankin & Baron 2005). And in yet another context, subjects were unwilling to pay any money at all to insure against a 1% chance of catastrophic loss (McClelland *et al.* 1993). So it seems that many of us, in many contexts, treat events with low probabilities as having probability 0, even when their true probability is as high as 1% (which is not so low). Upon reflection, this is clearly foolish. We may have the raw intuition that we may (or should) ignore all outcomes with probability 1% but it does not hold up to scrutiny, and so

[12] In effect, this is a core point of Bostrom's (2009) objection. Meanwhile, Bostrom (2011), Beckstead (2013), and Tarsney (n.d.) all treat Fanaticism as implausible for, it seems, the same reason.

is likely mistaken.

Given how widespread these mistakes are, we should give little weight to our intuitions about what we should do in cases of low probability, including those which lead us to recoil from fanatical verdicts - with a little more scrutiny, those intuitions may appear foolish too. At the very least, the intuitive case against Fanaticism is dubious enough that we should consider the case in favour of it.

3 Background assumptions

Before I make a case for Fanaticism, here are my basic assumptions.

First, for any decision problem, there is some set of epistemically possible outcomes $\mathcal{O}$. Some outcomes are better or worse than others. So assume that there exists some binary relation on $\mathcal{O}$, denoted by $\succcurlyeq_O$, which represents an 'at least as good as' relation between outcomes.

As mentioned earlier, for the most part I'll be assuming a totalist view of value. So $\succcurlyeq_O$ is already given: for any possible outcomes O_1 and O_2, we have $O_1 \succcurlyeq_O O_2$ if and only if O_1 contains at least as much total value as O_2. And total values are totally ordered, so we know that $\succcurlyeq_O$ will be reflexive, transitive, and complete on $\mathcal{O}$. These total values of outcomes can be represented with a cardinal value function $V : \mathcal{O} \to \mathbb{R}$, at least when those outcomes have finite differences in value.[13] And as a cardinal value function, V is unique only up to affine transformations - for any V_1 or V_2 we might use here, $V_2(O) = a \times V_1(O) + b$ for some positive a and real b.

We may not always be able to give a real representation of the total value for every outcome. If an outcome O_∞ contains infinitely more value than others - as when we create infinitely many blissful lives in Dyson's Wager - then it may fall beyond the scope of V. Then $V(O_\infty)$ won't be defined, but still $O_\infty \succ_O O$ for each finitely-valued outcome O. And that's fine - a real-valued total value function V won't represent $\succcurlyeq_O$ in all cases, but it will do just fine in purely finite cases. I'll use $\mathcal{O}_\mathbb{R} \subset \mathcal{O}$ to denote any given set of outcomes over which $\succcurlyeq_O$ admits a real-valued representation V. The arguments that follow can be run entirely in terms of $\mathcal{O}_\mathbb{R}$.

What we are really interested in is *lotteries*. I'll assume that the relevant aspects of a lottery can be fully described by a (real-valued) probability measure L on a set of outcomes. I'll also assume that this probability represents either the agent's subjective degree of confidence or the

[13] The differences in value for two pairs of outcomes are treated as finite precisely when one can be represented as a real multiple of the other.

evidential probability of each outcome arising. Almost all of what follows can be read in terms of either. With minor changes, what follows could also be read in terms of objective chance or whatever notion of probability you consider relevant for decision making. Whichever notion it is, let $\mathcal{L}$ be the set of all lotteries on $\mathcal{O}$, and $\mathcal{L}_{\mathbb{R}}$ the set of all lotteries on a given subset $\mathcal{O}_{\mathbb{R}}$. So each lottery L is a function which maps sets of possible outcomes to some value in the interval [0,1], and that function must obey the standard probability axioms.

To keep things brief, I'll use the following shorthand. When we are interested in the probability of a single outcome, I'll shorten $L(\{O\})$ to $L(O)$. When a lottery L gives a certainty of an outcome O (when $L(O) = 1$) I'll often write L as O. To represent a lottery with the same probabilities as L, but outcomes which have precisely k times the value (on the same cardinal representation), I'll write $k \cdot L$.[14] And I'll write $L_a + L_b$ to represent the lottery obtained from adding together the values of the outcomes of lotteries L_a and L_b, run independently.[15]

Fanatical verdicts, and their rejection, consist in *comparisons* of lotteries. So we need another 'at least as good relation'. Let $\succcurlyeq$ be a binary relation on $\mathcal{L}$ (and hence also on the subset $\mathcal{L}_{\mathbb{R}}$). Strict betterness ($\succ$) and equality ($\sim$) are defined as the asymmetric and symmetric components, respectively.

I won't assume that $\succcurlyeq$ is *complete* on $\mathcal{L}$. But I will assume that it is *reflexive*: that $L \succcurlyeq L$ for all $L \in \mathcal{L}$. I'll also assume that it is *transitive*: that, for all $L_a, L_b, L_c \in \mathcal{L}$, if $L_a \succcurlyeq L_b$ and $L_b \succcurlyeq L_c$, then $L_a \succcurlyeq L_c$. Both of these properties are highly plausible. If either of them do not hold, then betterness is a peculiar thing.[16]

I also want to assume another highly plausible principle of instrumental rationality, which will be useful in what follows: *Stochastic Dominance*. This principle says that if two lotteries have exactly the same probabilities of exactly the same (or equally good) outcomes, then they are equally good; and if you improve an outcome in either lottery, keeping the probabilities the same, then you improve that lottery. And that's hard to deny!

[14] Formally, for any real k and $L \in \mathcal{L}$, define $k \cdot L$ as a probability measure on $\mathcal{O}$ such that, for all O in $\mathcal{O}$, $k \cdot L(O_k) = L(O)$, where O_k is an outcome in $\mathcal{O}$ such that $V(O_k) = k \times V(O)$.

[15] This can be made more precise. Let V_i denote the random variable corresponding to a given lottery L_i, which outputs the total value of the outcome - equivalently, the probability that V_i takes on a value in the interval $[a, b]$ is given by $L_i(\{O|V(O) \in [a, b]\})$. Then, for any $L_a, L_b \in \mathcal{L}$, define $L_a + L_b$ as some lottery on $\mathcal{O}$ - there may be several - which corresponds to the random variable $(V_a + V_b)$. Equivalently, $L_a + L_b$ is some lottery such that $L_a + L_b(\{O|V(O) \in [c, d]\})$ is equal to the probability that $(V_a + V_b)$ takes on a value in the interval $[c, d]$.

[16] Some argue that moral betterness (and presumably also instrumental moral betterness) is not transitive, e.g. Temkin (2014). But the most compelling of these arguments assume pluralism with respect to moral value. But I'm focusing on monistic theories of value here, so transitivity remains a compelling principle.

We can express Stochastic Dominance as follows. Here, $\mathcal{O}_{\succcurlyeq O'}$ denotes the set of outcomes in $\mathcal{O}$ that are at least as good as O'.

> *Stochastic Dominance*: For any $L_a, L_b \in \mathcal{L}$, if $L_a(\mathcal{O}_{\succcurlyeq O'}) \geq L_b(\mathcal{O}_{\succcurlyeq O'})$ for all $O' \in \mathcal{O}$, then $L_a \succcurlyeq L_b$.
>
> If, as well, $L_a(\mathcal{O}_{\succcurlyeq O'}) > L_b(\mathcal{O}_{\succcurlyeq O'})$ for some $O' \in \mathcal{O}$, then $L_a \succ L_b$.

I find Stochastic Dominance overwhelmingly plausible. In this setting, it is far weaker than (but necessary for) expected value theory. But, unlike the stronger theory, it does not rule out risk aversion, nor Allais preferences. To deny it is to accept either: that you can swap an outcome in a lottery for a better one without making the lottery better; or that evaluations of lotteries are dependent on something other than the values of the outcomes and their probabilities. Both are implausible.[17]

And Stochastic Dominance in a setting like this is fairly uncontroversial among decision theorists, including some who reject expected value theory. As far as I know, no serious proposal has been made in normative decision theory which violates Stochastic Dominance (at least in this setting).[18]

And Stochastic Dominance ties together the fates of Fanaticism and the fanatical verdict in Dyson's Wager. Recall the lotteries L_{safe} and L_{risky} from the definition of Fanaticism. In L_{safe}, we can set $v = 1$. And define L_{infinite} as follows.

> L_{infinite}: an outcome containing infinitely many blissful lives with probability ϵ; value 0 otherwise
>
> L_{risky}: value V with probability ϵ; value 0 otherwise
>
> L_{safe}: value 1 with probability 1

L_{infinite} and L_{safe} should look familiar - these match the lotteries you must choose between in Dyson's Wager, if we represent the outcome in which one life is saved with value 1 and the outcome in which no lives are saved with value 0. By Fanaticism, L_{risky} is better than L_{safe}, at

[17] For compelling arguments in favour of Stochastic Dominance, see Easwaran (2014) and Bader (2018).

[18] Schoenfield (2014) rejects a similar principle, but only for lotteries over outcomes for which we don't have a complete ordering. With our complete $\succcurlyeq_O$ relation on $\mathcal{O}$, even Schoenfield would endorse Stochastic Dominance here.

least for large enough finite V. And Stochastic Dominance implies that L_{infinite} is better than L_{risky}.[19]

4 A continuum argument

You might argue for Fanaticism in the following way: expected value theory is true; expected value theory implies Fanaticism; therefore, Fanaticism is true. And likewise for verdicts which seem fanatical, like that which the expected value theorist must accept in Dyson's Wager. Such verdicts are rarely defended on any grounds other than 'That's what expected value theory says'. But we can do better than this. Fanaticism is far weaker than expected value theory in all its strength (and weaker even than expected *utility* theory), so it should be easier to justify.

In this and the next two sections, I'll give four arguments for Fanaticism but not for expected value theory. Here is the first, which originates in Beckstead (2013: 139-47) and reappears in Beckstead and Thomas (n.d.: 4-6).

Take the lottery L_0, which might represent a stranger's life being saved.

L_0: value 1 with probability 1

And take another lottery L_1 by which vastly more strangers are saved, with a very slight probability of failure - perhaps 10^{10} lives saved with probability 0.999999 of success.

L_1: value 10^{10} with probability 0.999999; value 0 otherwise

Intuitively, L_1 seems better. Accepting a *slightly* lower probability of success for a *vastly* greater payoff seems a good trade. But then consider L_2, which has a slightly lower probability of success but, if successful, results in many more lives saved.

L_2: value $10^{10^{10}}$ with probability 0.999999^2; value 0 otherwise

[19] The reasoning behind this is as follows. We represented the value of one additional life being saved with a 1 and additional lives saved with a 0. Any finite positive V will hence correspond to V additional lives of equal value being saved (or produced). This implies that the outcome in which infinitely many blissful lives are produced cannot be represented on the same scale, and indeed it is better than any outcome that can be. But, had that outcome been representable on the same scale with some finite number, this wouldn't hold.

This seems better than L_1, or at least it seems that there is some number of lives high enough that it would be better. And so we could continue, with L_3, L_4, and so on until some L_n, such that 0.999999^n is less than ϵ, for any arbitrarily small ϵ you want. Intuition suggests that *vastly* increasing the payoff can compensate for a *slightly* lower probability; that each lottery in the sequence is better than the one that precedes it. So the final lottery in the sequence must be better than the first. But the final lottery has a probability less than ϵ of any positive payoff at all. So we have Fanaticism.[20]

As I see it, this argument rests on two intuitively plausible principles: the first, the transitivity of $\succcurlyeq$; the second, *Minimal Tradeoffs*.

> *Minimal Tradeoffs*: There is some ratio $r < 1$ such that, for any real value v and any probability p, there is some real $r*$ such that, on any cardinal representation, the lottery represented by L_a is better than that represented by L_b (as defined below).
>
> L_a: value v with probability p; 0 otherwise
>
> L_b: value $r * \times v$ with probability $r \times p$; 0 otherwise

This principle has intuitive force. For any given lottery with the form of L_a, *surely* there is any lottery with only that ever-so-slightly lower probability of success is better. Intuition suggests that we can always make at least some tradeoff between probability and value: we can always compensate for a, perhaps, 0.000001% lower probability of success with a vastly greater payoff.

But you might be unconvinced by the continuum argument. You may find it far more intuitively plausible that Fanaticism is false than that Minimal Tradeoffs is true. And, other than clashing with intuition, I see nothing obviously problematic about denying Minimal Tradeoffs. Perhaps we should just reject it; perhaps there is at least one threshold of probability p' between p and 0 (which might be unknown or vague) at which we have a discontinuity, such that *no* value with probability below p' is better than *any* value with probability above p'. One way to do

[20] For readers less sympathetic to totalism, much the same argument can be made in terms of any of the axiologies mentioned in the Introduction. For instance, if we assume averageism, we can start with L_0, by which the *average* value of all lives will be 1 with near certainty, and 0 otherwise. L_1 can be the lottery which gives probability 0.99999 of an outcome in which the average value of all lives will be 10^{10}, or 0 otherwise. We can rapidly scale up the average value of lives just as we scaled up the total values above, while gradually scaling down the probability of non-zero value. Eventually, we reach a lottery L_n with probability ϵ of some astronomical average value, or 0 otherwise, but transitivity says that it must be better than L_0.

this is by adopting expected utility theory with a bounded utility function, which would imply that the former lottery is better than the latter. But, whichever approach we use to reach that conclusion, there is some point at which we can no longer trade off probability against value, no matter how great the value. And what of it? It seems a little counterintuitive, but perhaps less so than Fanaticism.

If you accept Minimal Tradeoffs then you must accept Fanaticism. And, inversely, to reject Fanaticism you must also reject Minimal Tradeoffs and the continuum argument it generates. But you might think that a small price to pay.

5 A dilemma for the unfanatical

The next argument for Fanaticism comes in the form of a nasty dilemma that you face if you deny it. In brief, if you deny Fanaticism then you must accept at least one of: abandoning *Scale Independence*, as defined below; and allowing comparisons of lotteries to be *absurdly sensitive* to tiny changes (even more so than if you accepted Fanaticism).

> *Scale Independence*: For any lotteries L_a and L_b, if $L_a \succcurlyeq L_b$, then $k \cdot L_a \succcurlyeq k \cdot L_b$ for any positive, real k.

I find Scale Independence highly plausible.[21] Take any pair of lotteries, perhaps a pair that can be cardinally represented by L_1 and L_2.[22]

L_1: value 1 with probability 1

L_2: value 2 with probability 0.9; value 0 otherwise

I'll remain agnostic on which is better; $\succcurlyeq$ might recommend either (or neither). But suppose we *scaled up* each lottery by 2, to obtain $2 \cdot L_1$ and $2 \cdot L_2$ - lotteries with the same probabilities

[21] One way to violate it is to adopt expected utility theory with a bounded utility function (see Rabin 2000). I find this good reason to reject such a form of expected utility theory. On the other hand, there are also theories that avoid Fanaticism without a bounded utility function and hence scale dependence, e.g., Buchak's (2013) risk-weighted expected utility theory and the Buffon-Smith method of ignoring outcomes with sufficiently low probabilities.

[22] For readers unsympathetic to totalism, these payoffs could just as easily represent: the average values of lives in each outcome; the priority-weighted total value in each outcome; the total value in the lives of all persons who would exist regardless of the agent's choice; and so on.

as above, but with the value of each outcome doubled (on the same cardinal scale). If we ranked L_1 as better then, intuitively, that ranking shouldn't change as we double the values. After all, the scaled-up lotteries have the same structure as the originals. Whatever general principles lead us to say whether L_1 is better than L_2 should plausibly also lead us to say the same about $k \cdot L_1$ and $k \cdot L_2$, for any positive k you like.

And remember that here we are dealing with moral value, not dollars or some other commodity that is merely *instrumentally* valuable. Unlike such commodities, when we are dealing with value itself, it is incoherent to say that additional units of value are worth less and less - by definition, adding one unit of value is always an improvement of one unit of value. And under totalism, some version of which I'm assuming here to be true, units of value can correspond to tangible objects such as (identical) human lives. By the lights of totalism and by at least my own intuition, an additional (identical) human life is always worth the same amount no matter how many lives already exist - and of course it always contributes precisely the same amount to the total value of the outcome. So, in keeping with intuition, it should not matter in the slightest whether we are comparing the lotteries L_1 and L_2 above or instead some scaled-up multiples $k \cdot L_1$ and $k \cdot L_2$. In the latter pair, we have done the equivalent of adding $k-1$ additional copies to the contents of each outcome (perhaps $k-1$ additional lives to L_1 and $2k-2$ additional lives to that outcome in L_2). But each additional copy should be no more or less valuable than the first; they should contribute the same to our evaluation of the lottery.[23] So, I would claim, we should rank the resulting lotteries just the same way as we ranked them without those $k-1$ copies added in.

It is also worth noting that totalism provides only that value may be represented cardinally, not that it can be represented on a ratio or unit scale (as do various other axiologies).[24] It does not provide any absolute zero of value, nor any unique scale on which to represent it. A pair of outcomes represented with values 1 and 2 on one scale can just as easily be represented with any values k and $2k$ (for positive k) on other scales. When evaluating those outcomes with totalism, we cannot say that any such representation is more valid than any other. Suppose we took

[23]A similar story can be told for averageism and other axiologies, although a little more awkwardly. On plausible versions of averageism, an additional period of bliss given to everyone contributes the same amount to the average value of all lives no matter how many such periods of bliss have already been experienced. We can copy the experiences each person has in much the same way we can copy the number of additional lives in the totalist scenario. It seems plausible then that we should rank lotteries with $k-1$ copies of each experience added into outcome just the same way as we rank them without those copies.

[24]Indeed, all of the axiologies listed in the introduction, with the exception of pure egalitarianism, typically do as well.

lotteries L_1 and L_2 above; their outcomes could just as easily have been represented on another scale such that they had values k, $2k$, and 0 - the same values as $k \cdot L_1$ and $k \cdot L_2$ had on the previous scale. On that new scale, we of course still say that L_1 is better than L_2 if and only if we said so on the original scale, else we would be inconsistent. But why not judge $k \cdot L_1$ and $k \cdot L_2$ similarly? For us to be able not to, our evaluations of lotteries must depend on more than just the probabilities of outcomes and cardinal values that totalism assigns them; they must also depend on *which* cardinal representation is being used when those values are assigned. In effect, they depend on richer details of the value of outcomes than is provided by the cardinal representation given by totalism. And this makes for a troubling shift in how we evaluate lotteries - doing so is no longer a matter of using our axiology to assign values to outcomes and then using our theory of instrumental rationality to turn those values and each outcome's probabilities into an evaluation. Instead, our evaluation is sensitive to facts about the values of outcomes that even our theory of value is not sensitive to. That is awfully strange and, I think, absurd. We should be able to apply our theory of rationality to the verdicts of our axiology without 'double-dipping' into facts about the value of outcomes. And so we should reject the suggestion that $k \cdot L_1$ and $k \cdot L_2$ not be compared in a similar way to L_1 and L_2; we should accept Scale Independence.[25]

For those who reject Fanaticism, the dilemma they face is between rejecting Scale Independence as well and accepting an absurd level of sensitivity to tiny changes in lotteries. I'll illustrate below what I mean by 'sensitivity to tiny changes' but, for now, be assured that it's just as implausible as violating Scale Independence.

The argument for the dilemma goes like this. First, for Fanaticism to be false, there must be some probability $\epsilon > 0$ and some cardinal value v such that some lottery L_{risky} as defined below is no better than L_{safe}, no matter how big V is.

L_{risky}: value V with probability ϵ; 0 otherwise

L_{safe}: value v with probability 1

But what about values below v? How would getting some lower value with certainty compare to an ϵ-probability of arbitrarily enormous V? There are only two distinct possibilities (at least

[25] An alternative to endorsing Scale Independence is to adopt a variant of totalism that assigns more than just a cardinal value to outcomes - perhaps a version of totalism that values outcomes on a ratio or unit scale. That version would be far less parsimonious than the version I have been using, and I cannot think of any compelling independent justification for it. It seems at least a considerable cost to adopt such a version of totalism to avoid endorsing Scale Independence and the implication of Fanaticism.

if we maintain Stochastic Dominance and transitivity). The first is: there is *some* smaller $v' > 0$ for which L_{risky} *is* better than a sure outcome with value of v', for sufficiently large V. The second is: for *any* such v', L_{risky} *isn't* better than any sure outcome with value v', no matter how large V is.[26] As we'll see, each possibility impales us on a respective horn of the dilemma.

Assume that the first possibility holds. L_{risky} is better than a sure outcome with value v', for some $0 < v' < v$ and some large V. Still, L_{risky} won't be better than a sure outcome with value v - i.e., L_{safe} - or any sure outcome with even greater value (by Stochastic Dominance). But somewhere below v this changes - for some positive v', the same doesn't hold.

Since that v' is a positive real number, we know that there will be some real k such that $k \times v' \geq v$. And that means that a sure outcome with value $k \times v'$ is no worse than L_1, no matter how great V is. But here is the problem: a certainty of $k \times v'$ is the scaled-up counterpart of v'; and the counterpart of L_{risky}, which is $k \cdot L_{\text{risky}}$), is just L_{risky} with V replaced by another value which is k times higher. That still has the same form as L_{risky}. So it cannot be better than the sure outcome, not if we reject Fanaticism. But then our verdict for v' versus L_{risky} doesn't match our verdict for $k \times v'$ versus $k \cdot L_{\text{risky}}$. So we violate Scale Independence.

But there was a second possibility: that there is no such $v' > 0$; that L_{risky} is no better than a certainty of *any* $v' > 0$. If so, we can avoid the problem of scale dependence: there's no inconsistency between the judgements for v and judgements for any 'scaled-down' counterpart v', since *all* such v' compare the same way to lotteries like L_{risky}.[27] But we then face another serious problem.

Take the probabilities ϵ and 1. We can give a sequence of increasing probabilities in between ϵ and 1, spaced evenly apart and as finely as we want: $\epsilon < p_1 < p_2 < \ldots < p_n < 1$. By assumption, *no* amount of value with probability ϵ (and value 0 otherwise) is better than *any* (positive) amount of value with probability 1. But then there must be some pair of successive p_i, p_{i+1} such that no amount of value at p_i is better than any amount of value with probability p_{i+1}. If there were no such pair, we could give a sequence of better and better, and less and less likely, lotteries much like that in the previous section.

Crucially, that p_i and p_{i+1} can be *arbitrarily* close together, since the same result holds no matter how finely spaced the sequence was. No amount of value at p_i is better than *any* amount

[26] This difference is analogous to the difference between weak and strong superiority in Arrhenius & Rabinowicz (2015).

[27] The comparison won't be the same for values 0 and below, but that's okay. Such values cannot be scaled up to v; multiply 0 or any negative number by any real positive value you want and you still won't reach v.

of value at p_{i+1}. We could have some *astronomical* value at p_i, and an imperceptibly small amount of value at p_{i+1}, and the latter would still be better. So we have a radical discontinuity in how we value lotteries.

This will make our judgements of betterness *absurdly sensitive* to tiny changes in probability. To illustrate that sensitivity, consider the following four lotteries. Here, $\epsilon' > 0$ is some arbitrarily small number. And, again, p_i and p_{i+1} are some tiny probabilities that are arbitrarily close together.

L_0: value 0 with probability 1

L_1: value $10^{10^{10}}$ with probability p_i; 0 otherwise

L_2: value ϵ' with probability p_{i+1}; 0 otherwise

L_3: value $10^{10^{10}}$ with probability p_{i+1}; 0 otherwise

By the above, if we reject Fanaticism and maintain Scale Independence, then we are forced to rank these lotteries as follows: $L_0 \prec L_1 \prec L_3$ and $L_0 \prec L_2 \prec L_3$, but $L_1 \not\prec L_2$. And that's a peculiar ranking. L_1 and L_3 are almost indistinguishable; their probabilities may differ by an arbitrarily tiny amount. Likewise for L_0 and L_2, except it's their payoffs that differ slightly. For each pair, we must say that one is better than the other, but does not seem *much* better - in a sense, the better lottery is only a trivial improvement over the other, whether by a slight increase in payoff or in probability. But, despite appearances, we must accept that L_3 is *much* better than L_1, as well as that L_2 is much better than L_0. How so? Between L_1 and L_3 comes L_2. The lottery L_2 has an *astronomically* smaller payoff than L_3 and so (in a sense) is *vastly* worse than L_3, and yet we cannot say that it is so bad so as to be worse than L_1. Effectively, we must accept that, starting from L_3, it is no worse to lose almost all of the potential value via L_2 than it is to lose that sliver of probability via L_1. And similarly for L_0 and L_2: between them must come L_1. L_1 has astronomically greater potential payoff than L_0 (and much higher probability of positive payoff too), yet we cannot say it's better than L_2. Effectively, we cannot say it is any better to gain that enormous value with probability p_i than it is to gain an ever-so-slightly more probable shot at tiny value ϵ. So those lotteries are not just made worse by those tiny changes in probability or payoffs - they are made *far* worse (in an intuitive sense). That is the level of sensitivity we must accept if we reject Fanaticism and maintain Scale Independence. But such extreme sensitivity in our evaluations of lotteries seems absurd. Evaluations of lotteries should not change quite so wildly - so discontinuously - with arbitrarily tiny changes in probability or payoff. So I do not think this sensitivity - this horn of the dilemma - is any less absurd than what we faced earlier.

I should briefly note that, beyond its intuitive implausibility, this sensitivity will likely lead to practical difficulties. The probabilities that human agents have access to are subjective degrees of belief and evidential probabilities. In practice, neither sort of probability can be pinned down with arbitrary precision, at least not with our merely finite capacity for calculation.[28] But we'll sometimes need arbitrary precision to compare lotteries, at least by this horn of the dilemma. If not, in cases like that immediately above, we would have no idea whether a lottery is as good as L_3 or as bad as L_1 with its imperceptibly lower probability of success. We must either require (unrealistically) that agents be arbitrarily precise in their probabilities[29], or else accept that in practice our decision theory will be unable to tell agents that (at least some of) these lotteries are better than others.

This is a dilemma we must face if we reject Fanaticism: we must accept either scale dependence, by which our judgements of lotteries can vary without any structural reasons for doing so, or that those judgements are sensitive to imperceptibly small differences in either probability and value. Both options seem absurd to me, indeed more absurd than Fanaticism itself.

6 Egyptology and Indology

But it gets worse. There are even greater costs you must pay to reject Fanaticism: you face either one or both of the following objections, each of which is intuitively absurd.

6.1 The Egyptology Objection

Hailing from population ethics, the *Egyptology Objection* is a classic argument against various axiologies, including averageism, egalitarianism, and maximin.[30]

[28]For discussion of why we cannot require epistemic agents to settle on precise probabilities, see Schoenfield (2012) and Joyce (2011). For discussion of how to evaluate lotteries with imprecise probabilities, see Seidenfeld (2004), Huntley, Hable & Troffaes (2014), Bradley & Steele (2015), and Bradley (2015). On any of the methods suggested, a combination of imprecise probabilities and sensitivity to the degree described above will lead to indifference among many of the lotteries L_1 to L_4 listed above or else indeterminacy of which is better than which. At least in that case, with its vast differences in value and only small differences in probability, it ***

[29]In Smith's (2014:471) terms, our theory would then be 'intolerant' of human imprecision, which he considers a major problem.

[30]It appears earliest in MacMahan (1981: 115) but is often attributed to Parfit (1984: 420).

As a brief refresher, it goes like this. Suppose you are making a moral decision here and now in the 21st Century, and the available outcomes differ *only* by some small-scale changes in the very near future. Your actions will not change what happens in distant galaxies or what happened in, say, ancient Egypt. Then, intuitively, what you ought to do should depend only on the events altered by your choice. It should not depend on events in distant galaxies or in ancient Egypt, at least not if your actions do not change them. Intuitively, it seems absurd that those unaltered, remote events make any difference to your evaluation.[31]

But, under some axiologies, what you ought to do in such cases *is* dependent on unaltered, remote events. Take a (standard, welfarist) averageist view. Here and now, should you bring an additional person into existence? It depends - will that additional life raise or lower the average value of lives that ever exist? Sometimes yes, sometimes no, depending on what actually happened in ancient Egypt, in distant galaxies, and everywhere else. Averageism will thus imply that what happened in ancient Egypt can affect whether it is better to now bring an additional person into existence or not. And, further, to be confident of which is the better outcome, you may need to do some serious research into Egyptology (along with plenty of other historical studies). But this is implausible. By intuition, we can ignore events that are unaffected by our actions and took place millennia ago, as we do under axiologies like totalism, critical-level views, prioritarianism, and aggregative person-affecting views (and any other axiology which admits an additively separable representation). To me, this is one of the main appeals of those views.

But even if we accept an axiology like totalism, we may still face much the same objection in practice.[32] Some methods of comparing lotteries give rise to an updated Egyptology Objection, including all of those that deny Fanaticism. That's right: to deny Fanaticism, you must not only endure the costs detailed in the previous two sections, you must also fall prey to a version of the Egyptology Objection. (As we'll see below, you may also face another even more serious objection.)

You will most easily fall prey to it if you reject *Background Independence*. (And some proposals do.)[33]

[31]Such dependence brings practical problems too. If what you ought to do is dependent on such events - not just in the remote future but also the remote past, of which we are often clueless - then we have little guidance for what we ought to do (*cf* Mogensen 2021). This seems a major disadvantage for a moral view.

[32]Even on other axiologies, including averageism, we face similar (albeit slightly less compelling) objections. See Footnote 35.

[33]One such proposal is expected utility theory with a utility function that is concave and/or bounded (e.g., Arrow 1971). As Beckstead and Thomas (n.d.: 15-16) point out, this results in comparisons of lotteries being strangely dependent on events that are unaltered in every outcome and indeed some irrelevant to the comparison.

Background Independence: For any lotteries L_a and L_b and any outcome O', if $L_a \succcurlyeq L_b$, then $L_a + O' \succcurlyeq L_b + O'$.

Recall that the sum of two lotteries $L+O'$ is simply the lottery you get if you run each lottery independently and sum up the values of their outcomes. But, here, O' is a very simple lottery - it is just the outcome O' with certainty. That outcome has some cardinal value; call it b. The lottery $L_a + O'$ is simply the lottery L_a with value b added to the value of each outcome. So Background Independence says that we can take any two lotteries and add a constant value to every one of their possible outcomes, and this won't change their ranking.

Suppose for now that Background Independence is false. Then there is some pair of lotteries, L_a and L_b, such that: L_a is at least as good as L_b; and, if we added a certain constant b to the value of every one of their outcomes, that would change their ranking. Since we're assuming totalism, the value of each outcome in each lottery must represent the total aggregate of value that would result, including all valuable events across all of space and time. That includes the events that occurred in ancient Egypt. And we can assume that the same events occurred in ancient Egypt in every outcome of either L_a or L_b.[34] With those events included, L_a is at least as good as L_b.

But what if we consider a different hypothetical pair of lotteries, identical to L_a and L_b *except* that the events of ancient Egypt were different? Would the ranking be any different, had events gone differently back then? Yes, it would, if Background Independence is false and if those events differed drastically enough. They might differ in such a way that they increased in value by b. Then the total value of every outcome would also be increased by b. And, given that L_a and L_b were the lotteries that violated Background Independence, we know that this changes their ranking. So our modified version of L_a is no longer at least as good as the modified version of L_b. Thus, if we deny Background Independence, we face a rather severe version of the Egyptology Objection: when choosing between two lotteries, which is better can vary depending on what events occurred in ancient Egypt, *even if* the choice doesn't change those events at all (*cf* Beckstead and Thomas n.d.).[35]

[34] If there is *some* pair of lotteries L_a or L_b that violate Background Independence, then there is also some pair that violate it and have the exact same events occur in ancient Egypt. To obtain such a pair, simply take L_a and L_b and replace each of their outcomes with an outcome of equal total value but an identical history of ancient Egypt.

[35] This new version of the Egyptology Objection, along with the other objections described in the rest of this section, can also be applied under axiologies other than totalism. It will just look a little different.

Suppose that (standard, welfarist) averageism is the correct axiology. Then to add b to value of an outcome

To avoid this implication, we must at minimum accept Background Independence. I think we ought to accept it regardless - I find it highly plausible for much the same reasons as Scale Independence was. In any case, I'll assume for the remainder of this section that Background Independence holds.

But even if we accept Background Independence, if we still reject *Fanaticism* then it turns out that a version of the Egyptology Objection still arises, albeit a less severe one.

To see why, first recall the lotteries from the definition of Fanaticism: L_{risky} and L_{safe}. If we reject Fanaticism then, for some tiny enough probability $\epsilon > 0$ and some cardinal value v, some lottery L_{risky} (as defined below) is no better than L_{risky}, no matter how big V is.

L_{risky}: value V with probability ϵ; 0 otherwise

L_{safe}: value v with probability 1

According to Background Independence, it must also hold that the lottery L_{risky} with any constant b added to every outcome's value is also no better than L_{safe} with that same b added to every one of its outcomes' values. If we face a decision between L_{risky} and L_{safe}, it would not matter if events in ancient Egypt were vastly different - we'd compare the lotteries the same way with or without that additional value b.

But consider a further pair of lotteries: $L_{\text{risky}} + B$ and $L_{\text{safe}} + B$. These are obtained by modifying L_{risky} and L_{safe} such that, in all outcomes, the events of ancient Egypt are different but you are *uncertain* of exactly how different they are. We are no longer dealing with a fixed value b of which you are certain. We are still adding the same value to all of them, but you are uncertain of what value it is. And that uncertainty is described by the lottery B. To make these lotteries more concrete, perhaps you must decide between two lotteries in which the value of all present and future events match the lotteries L_{risky} and L_{safe}, but you are uncertain of the value of past events such as those in ancient Egypt. As a devout totalist, you must evaluate outcomes

would be to increase the average value obtained by all persons in that outcome. And to add b to the value of every outcome in a lottery L would be to increase the average value in every outcome by b. We might imagine doing this by delivering a gift to every inhabitant of the outcomes of L, with that gift producing the same boost in value, b, for each recipient. Even by the lights of averageism it seems that, if L_a is at least as good as L_b, then delivering that gift to everyone shouldn't change the ranking - a modified L_a with additional value b should still be at least as good as L_b with additional value b. And so we will have an averageist analogue of the Egyptology Objection presented above. And this objection will still be worrying for averageists, even if it is not quite as devastating as its analogue is for totalists.

and lotteries based on the total aggregate of value across all of space and time. And even if you know that those past events will turn out the same way no matter what you do, your uncertainty over exactly how they would turn out is a part of your uncertainty over the total value of the outcome. So you must compare $L_{\text{risky}} + B$ and $L_{\text{safe}} + B$; you cannot simply L_{risky} and L_{safe}.

As it turns out, there are some possible lotteries B that lead us to say that $L_{\text{risky}} + B$ is strictly better than $L_{\text{safe}} + B$, even though L_{risky} isn't any better than L_{safe}. And this judgement, that $L_{\text{risky}} + B$ is better, is implied by even the extremely weak principle of Stochastic Dominance introduced above.[36]

As a brief refresher, recall that Stochastic Dominance simply states that: if a lottery L_a gives at least as high a probability as L_b of resulting in an outcome which is at least as good as O, for *every* possible outcome O, then L_a is at least as good a lottery; and if L_a also gives a *strictly greater* probability of turning out at least as good as some O, then it is *strictly* better.

Stochastic dominance is easy to spot graphically. To illustrate, consider the cumulative probability graphs of the following two lotteries, L_1 and L_2.

L_1: value 1 with probability $\frac{1}{2}$; value 0 otherwise

L_2: value 2 with probability $\frac{1}{3}$; value 1 with probability $\frac{1}{3}$; value 0 otherwise

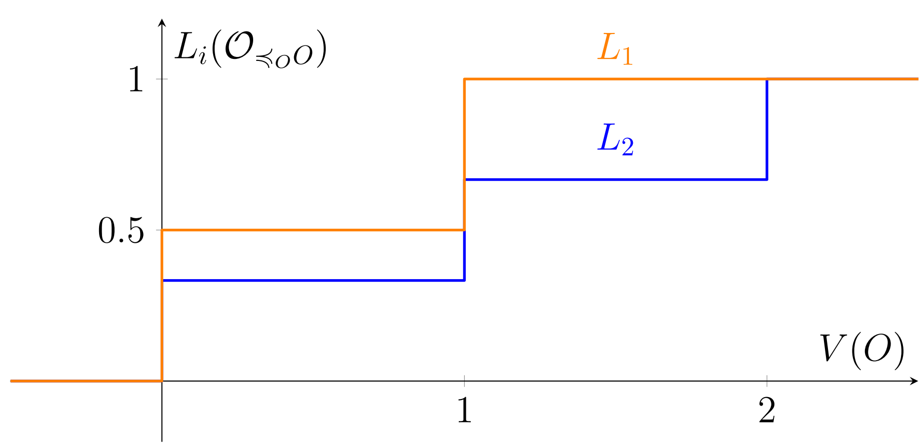

Figure 1: Cumulative probability graphs for L_1 and L_2; here and below, $\mathcal{O}_{\preccurlyeq_O O}$ denotes the set of outcomes in $\mathcal{O}$ as which O is at least as good.

[36] We can reach a similar result, and thus a similar *reductio*, with axiologies other than totalism. Following from Footnote 35, we can adopt averageism and, instead of improving/worsening events in ancient Egypt, distributing an identical and morally valuable gift to each person in the world. You might be uncertain of the exact value of the gift - your uncertainty of its value might correspond to B. But still it seems that whether you distribute that gift is irrelevant to whether the risky lottery is better or worse than the safe one.

On a graph like this, one can easily see when Stochastic Dominance says that $L_2 \succcurlyeq L_1$. Cumulative probability, on the vertical axis, is just the probability that the lottery produces an outcome *no better* than an outcome with some particular value. Meanwhile, Stochastic Dominance says that one lottery is at least as good as another if its probability of producing an outcome *as good or better* is just as high for all outcomes - or, equivalently, if the probability of an outcome *as good or worse* is at least as low. So Stochastic Dominance will say that $L_2 \succcurlyeq L_1$ if and only if L_2's cumulative probability is always as low or lower than that of L_1, as it is on this graph. And, here, L_2 often has *strictly* lower cumulative probability. So Stochastic Dominance says it is strictly better than L_1.[37]

But accepting Stochastic Dominance doesn't rule out denying Fanaticism. L_{risky} doesn't stochastically dominate L_{safe}, as illustrated below.

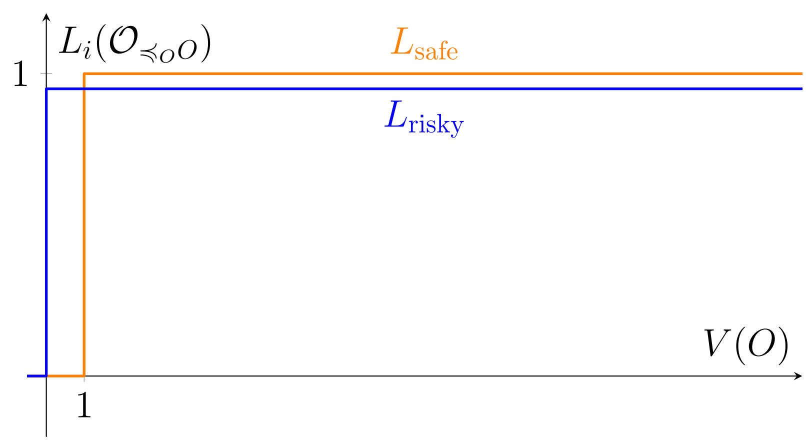

Figure 2: Cumulative probability graphs for L_{risky} and L_{safe}

Sometimes one is higher; sometimes the other. So Stochastic Dominance remains silent. And it's a good thing it does - to deny Fanaticism, we must not say that L_{safe} is better than L_{risky}.

But what about $L_{\text{risky}} + B$ and $L_{\text{safe}} + B$? Stochastic Dominance is not necessarily silent about that comparison. In particular, suppose that the lottery B looks like this. (I'll describe the required properties of B below.)

[37]This relationship between Stochastic Dominance and cumulative probability would break down if some outcomes were incomparable to others. There would then be a difference between $O_a \succcurlyeq_O O_b$ and the negation of $O_a \prec_O O_b$. But, fortunately, totalism gives us a total preorder of $\mathcal{O}$, so we can sidestep this complexity.

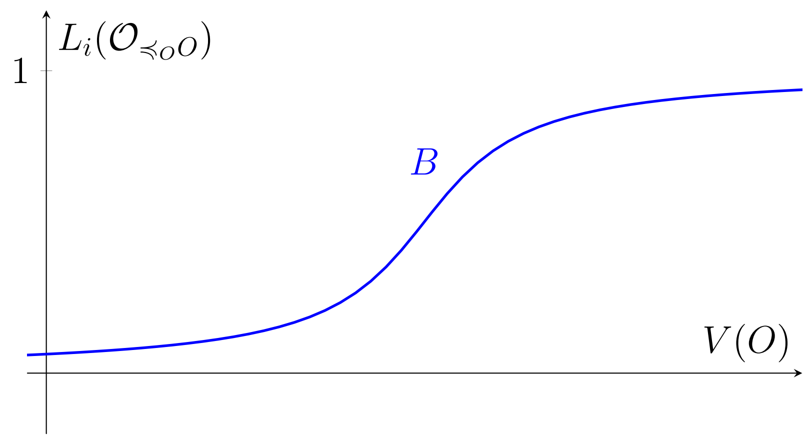

Figure 3: An example of background uncertainty B - a Cauchy distribution

When B looks like that, we obtain the following graphs for $L_{\text{risky}} + B$ and $L_{\text{safe}} + B$. Crucially, the graph for $L_{\text{safe}} + B$ is never higher than that for $L_{\text{risky}} + B$, and sometimes it is strictly lower. So Stochastic Dominance says that $L_{\text{safe}} + B$ is strictly better.

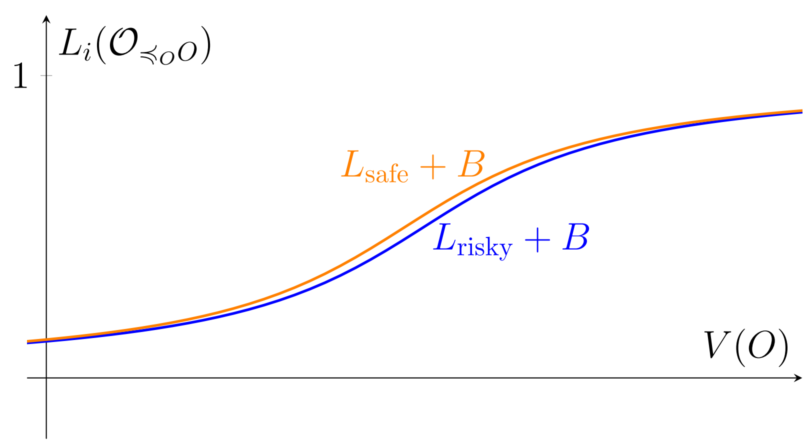

Figure 4: Cumulative probability graphs for $L_{\text{risky}} + B$ and $L_{\text{safe}} + B$

But how does this happen?

For Stochastic Dominance to hold, we need $L_{\text{risky}} + B$ to have at least as high a probability as $L_{\text{safe}} + B$ of turning out *at least as good as* an outcome with value u, for *all* possible values u. So take any real value $u < V$. What's each lottery's probability of doing at least that well?

Start with $L_{\text{risky}} + B$. We know that L_{safe} just gives one value (v) with certainty. So the probability that $L_{\text{safe}} + B$ gives value u or better is just the probability that B gives value $u - v$ or better. (This corresponds to the area $B_2 + B_3$ on the graph below.) And then, for $L_{\text{risky}} + B$, we'll get value at least u either if L_{risky} gives value V or if B gives at least value u. Denote the probability that B gives value at least u by B_3 (corresponding to that area on the graph below). Then the probability that $L_{\text{risky}} + B$ turns out at least as good as value u is $\epsilon + B_3(1-\epsilon)$. And, with a bit of simple arithmetic[38], we can see that this will be greater than the corresponding probability for $L_{\text{safe}} + B$ if the area B_2 is no greater than $\epsilon \times B_1$.

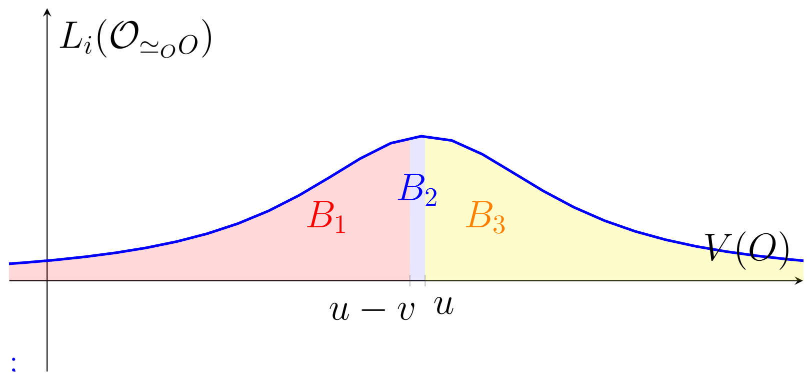

Figure 5: Probability distribution of B

Here we have a probability graph of one possible B (importantly, not a *cumulative* probability graph), with areas B_1, B_2, and B_3 corresponding to the areas mentioned in the last paragraph. If B_2 is small enough compared to B_1, then $L_{\text{risky}} + B$ has at least as high a probability of u or better as $L_{\text{safe}} + B$ does. And for it to be at least as high for *all* real u, we just need the interval between u and $u - v$ to have a tiny enough area under it. And, for that to happen, we just need B to go down slowly enough as we approach $-\infty$, and rise and fall quickly enough as we pass the peak of the curve.

And there are some probability distributions that have this property, including many Cauchy distributions.[39] And some is enough. There is some such lottery $L_{\text{risky}} + B$ that is better than $L_{\text{safe}} + B$, even though L_{risky} is no better than L_{safe}.

And so we have a new version of the Egyptology Objection, which I think is slightly milder than the previous version. You may be faced with a decision between two lotteries that, if you

[38]
$$\begin{aligned} \epsilon + B_3(1-\epsilon) \geq B_2 + B_3 &\Leftrightarrow B_3 + \epsilon(1 - B_3) \geq B_2 + B_3 \\ &\Leftrightarrow \epsilon(B_1 + B_2) \geq B_2 \\ &\Leftrightarrow B_2 \leq \epsilon B_1 + \epsilon B_2, \qquad \text{for which } B_2 \leq \epsilon B_1 \quad \text{is a sufficient condition.} \end{aligned}$$

[39]The result holds for any Cauchy distribution with a 'scale factor' of at least $\frac{v}{\epsilon}$. See Tarsney (n.d.: §5) and Pomatto *et al* (2018).

exclude the value of events that occurred in ancient Egypt, correspond to L_{risky} and L_{safe}. And which lottery is better overall? That depends on what happened in ancient Egypt, even though you know the same events will have happened there no matter which you choose. If the events of ancient Egypt have value 0 (or b), then the risky lottery is no better than the safe one. But if ancient Egypt might have contained greater or lesser value, and you aren't *certain* how much, then it may be that the risky lottery is better after all. Much like under the classic Egyptology Objection, your evaluation is sensitive to facts that seem irrelevant. But, unlike the original version, the evaluation is not merely sensitive to what *actually* occurred there; instead, it is sensitive just to your *uncertainty* about what happened there. That is perhaps less devastating a problem, as an agent may know enough about events in ancient Egypt to constrain B to a less troublesome distribution. Still, for judgements to be sensitive to the agent's beliefs about events in ancient Egypt *at all* is still, I think, rather absurd - absurd enough that we should be willing to accept Fanaticism to avoid it.[40]

6.2 The Indology Objection

But it gets worse. If you reject Fanaticism, you may also face what I'll call the *Indology Objection*.

For context, Indology is the study of the history and culture of India. In particular, classical Indology includes the study of the Indus Valley Civilisation, a Bronze Age civilisation which lasted from 3300 BCE to 1300 BCE. Althought less well-known and understood than its contemporaries - ancient Egypt and Mesopotamia - it spanned a greater area than either and rivalled them in population, technology, and urban infrastructure (see Wright 2009). We happen to know even less about what happened in the ancient Indus Valley than in ancient Egypt - archaeological research and excavations of key sites in India began centuries later than similar work in Britain, Italy, and Egypt. So there is likely plenty left to learn in Indology.

Take the same $L_{\text{risky}} + B$ and $L_{\text{safe}} + B$ from above - lotteries such that the former is strictly better, and yet L_{risky} is no better than L_{safe}. Such lotteries must exist, if Fanaticism is false. But

[40]This version of the Egyptology Objection loosely resembles one which has been levelled against decision theories that violate the Sure-Thing Principle (by, e.g., Briggs 2015). It turns out that following such a decision theory in each of several consecutive, independent decisions will sometimes be inconsistent with following the theory when deciding on a strategy for all of those decisions at once. Essentially, the combination of two lotteries can be evaluated quite differently to how we evaluate those lotteries separately. And that is troubling. It is unclear whether we should evaluate such decisions together or apart, globally or individually. But the correct verdict depends on which way we do it.

suppose that you are no longer uncertain about the events of ancient Egypt; instead, you are uncertain of what occurred in the ancient Indus Valley. The lottery B in $L_{\text{risky}} + B$ and $L_{\text{safe}} + B$ represents only those events.

But, in Indology, there is a great deal more work that can be done! Had you the option, you could spend many years researching the ancient Indus Valley, peeling back that uncertainty and narrowing B down. Let's suppose that with enough years of intensive research you could eventually remove that uncertainty entirely and determine exactly which value to assign to the events of the ancient Indus Valley - which value b the lottery B actually results in. If you did first do that research, you would no longer need to choose between $L_{\text{risky}} + B$ and $L_{\text{safe}} + B$. Instead, you would choose between L_{risky} with b added to the value of every outcome and L_{safe} with b added.

Recall Background Independence from above. If it holds, then L_{risky} with any b added to every outcome is better than L_{safe} with the same b added if and *only if* L_{risky} is better than L_{safe}. And if it doesn't, as we saw above, we face an even more severe version of the Egyptology Objection. So we can safely assume that Background Independence holds.

From Background Independence we know that, whatever you might uncover in your research, you would conclude that the risky lottery is no better than the safe lottery. For *any* value of b you might pin down, you'll establish that it's no better to take the risky lottery. To judge otherwise would be to accept Fanaticism. So you know what judgement you would make if you simply learned more, *no matter what* it is you would actually learn. So why do the many years of research? Why not just update your judgement now and say that $L_{\text{risky}} + B$ is no better than $L_{\text{safe}} + B$? You cannot. You must accept that $L_{\text{risky}} + B$ is strictly better, even though you can predict with certainty that you would change your mind if you knew more. To be able to change your mind you must go through those years of gruelling research and digging, even though you already know which judgement you would conclude to be correct.[41]

This ambivalence seems deeply irrational. Surely we can sidestep those years of research into how B turns out, and make the judgement required by every possible value of b. Surely rationality requires that we do so, rather than require that we do not. But, if we deny Fanaticism, we must accept this inconsistency - an inconsistency which, to me, seems far more absurd than

[41] *cf* van Fraassen's (1984) principle of *Reflection* and its application by White (2010:17-8). That form of Reflection implies that, if you know that you would set the probability of some hypothesis to p whether you learned either proposition P or not-P, then its current probability should already be p. The version I am assuming is somewhat similar but need not be quite as strong.

simply accepting Fanaticism and even more absurd than the Egyptology Objection.[42] [43]

Taking stock of this section, we have further reasons to accept Fanaticism. If we deny it then we are forced to accept a version of the Egyptology Objection wherein the judgements of agents like us will be sensitive to their uncertainty about what occurred in ancient Egypt, even though their choices do not affect ancient Egypt at all. If we reject both Fanaticism and Background Independence, then we face an even more severe version of the Egyptology Objection: how such agents should compare their options will be sensitive to not just their uncertainty but also what actually occurred in ancient Egypt. And even if we accept Background Independence (as I think we should), if we try to reject Fanaticism then a bizarre form of inconsistency arises in how agents should compare their options - agents are sometimes forced to make judgements which are inconsistent with *anything* they might learn to resolve their uncertainty about far-off, unrelated events in the ancient Indus Valley. Any one of these implications - let alone any two of them - is absurd. In light of these, it seems a lot less appealing to deny Fanaticism.

[42] Continuing from Footnotes 35 and 36, an analogue of the Indology Objection applies to averageism. (Analogues can also be constructed for other axiologies.) You might be facing two lotteries - one risky and one safe - in which you know that, whatever the outcome, an identical gift will be given to everyone in the world. And you're uncertain of just how much value that gift will bring, with your uncertainty corresponding to B. Then, as we saw above, there are some probability distributions for B that mean that you will judge $L_{\text{risky}} + B$ as strictly better than $L_{\text{safe}} + B$; but, if you opened the gift yourself and discovered what was inside, you already know that L_{risky} with the additional value b would be no better than L_{safe} with the additional value b. And this seems absurd, just as its analogue was for the totalist.

[43] A closely related objection to rejecting Fanaticism is that, even if we do reject it, we will not be spared from many verdicts that seem deeply fanatical. As we saw in the case above, with the right uncertainty B about the value of unaffected events you must judge $L_{\text{risky}} + B$ better than $L_{\text{safe}} + B$. In practice, will we often have the right uncertainty about unaffected events to lead us to judge otherwise extremely risky lotteries as better than otherwise safe ones? Tarnsey (n.d.: §6) argues that we often will, for an epistemically rational agent and for a wide range of cases, including many resembling L_{risky} and L_{safe} (where the probability ϵ is at least 1 in 10^9, and likely also a lot lower). His reasoning for this is that the moral value of distant events in our universe will be roughly correlated with the number of inhabited planets in the universe; and our cosmological evidence and the Drake equation together motivate a probability distribution over that number of inhabited planets which is relevantly similar to the above B. But even apart from that reasoning, an epistemically rational agent would still have some uncertainty over which probability distribution best characterises our cosmological evidence. So they would still put some non-zero credence in distributions of the right sort. And this is enough - when we take the probability-weighted average of many such distributions to get our *overall* distribution, it will then have the right properties (see *ibid.*: 37-9 for details). Given this, Stochastic Dominance by itself will lead us to making judgements which seem fanatical. So we would gain little from rejecting Fanaticism.

7 Conclusion

These are just some of the compelling arguments for accepting Fanaticism and, with it, fanatical verdicts in cases like Dyson's Wager. There are other compelling arguments too - for instance, all of the arguments for the stronger claims of expected value theory, or for expected utility theory with an unbounded utility function. *A fortiori*, these also justify Fanaticism. For a sampling of these, I refer interested readers to Feller (1968), Ng (2012), Briggs (2015), and Thoma (2018).

Some philosophers still reject expected value theory (e.g., Buchak 2013, Smith 2014, Monton 2019, Tarsney n.d.), and presumably the arguments that imply it. But this is not enough to escape Fanaticism. As I have demonstrated here, there are compelling arguments in its favour - arguments stronger than many of those for expected value theory.

To recap, if we deny Fanaticism then we must deny either Minimal Tradeoffs or transitivity, and accept the counterintuitive verdicts which follow. Likewise, we must accept either: violations of Scale Independence; or absurd sensitivity to the tiniest differences in probabilities and value. So too, to deny Fanaticism, we must accept a version of the Egyptology Objection: for agents like ourselves, which judgement is correct can depend on our beliefs about what occurred in far-off, unrelated events, such as those in ancient Egypt. And, unless we wish to accept an even more severe version of the Egyptology Objection, we also face the *Indology* Objection: we sometimes ought to make judgements that we know we would reject if we learnt more, no matter *what* we might learn.

Given all of this, it no longer seems so attractive to reject Fanaticism. As it turns out, the cure is worse than the disease. I would suggest that it is better to simply accept Fanaticism and, with it, fanatical verdicts such as that in Dyson's Wager. We should accept that it is better to produce some tiny probability of infinite moral gain (or arbitrarily high gain), no matter how tiny the probability, than it is to produce some modest finite gain with certainty.

All of this has implications for normative decision theory more broadly, at least dialectically. Philosophers often reject expected value theory because it implies Fanaticism, or because it implies fanatical verdicts in cases like Dyson's Wager (e.g., Monton 2019, Tarsney 2020). The arguments I have given here suggest that this rejection is a little hasty. Doing so invites greater problems than it solves. So I, for one, am thankful that expected value theory *does* imply Fanaticism. We have little choice but to accept Fanaticism, so we might as well accept expected value theory too.[44]

[44]This paper has benefitted from the input of many brilliant and generous colleagues. I am indebted to Alan

References

[1] Arrhenius, Gustaf 2000. *Future Generations: A Challenge for Moral Theory* (Uppsala: University Press).

[2] Arrhenius, Gustaf, & Wlodek Rabinowicz 2015. 'Value superiority', *The Oxford Handbook of Value Theory*: 225-48.

[3] Arrow, Kenneth 1971. *Essays in the Theory of Risk-Bearing* (Chicago: Markham).

[4] Askell, Amanda 2019. *Pareto Principles in Infinite Ethics*, PhD dissertation, New York University.

[5] Bader, Ralf 2018, 'Stochastic dominance and opaque sweetening', *Australasian Journal of Philosophy* 96, no. 3:498-507.

[6] Baron, Jonathan, Laura Granato, Mark Spranca, & Eva Teubal 1993. 'Decision making biases in children and early adolescents: Exploratory studies', *Merrill Palmer Quarterly* 39: 23–47.

[7] Beckstead, Nicholas 2013. *On the Overwhelming Importance of Shaping the Far Future*, PhD dissertation, Rutgers University.

[8] Beckstead, Nicholas and Teruji Thomas. 'A paradox for tiny probabilities and enormous values', unpublished manuscript.

[9] Bostrom, Nick 2009. 'Pascal's mugging', *Analysis* 69, no. 3: 443-5.

[10] Bostrom, Nick 2011, 'Infinite ethics', *Analysis and Metaphysics* 10:9.

[11] Buchak, Lara 2013, *Risk and Rationality* (Oxford: Oxford University Press).

[12] Bradley, Seamus 2015. 'How to choose among choice functions', *Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications*: 57–66.

Hájek, Christian Tarsney, Teru Thomas, Jasper Hedges, Timothy L. Williamson, Andreas Mogensen, and two anonymous reviewers for detailed comments. Likewise to Philip Trammell, Anders Sandberg, Shang Long Yeo and audiences at both the Australian National University and the Forethought Foundation for valuable discussions. For their generous financial support, I thank the Australian Government Research Training Program and the Forethought Foundation for Global Priorities Research. As well, I am grateful to all of the staff at the Global Priorities Institute at the University of Oxford for inviting me to visit for several months in late 2019 and early 2020, and to convene a reading group on infinitary fanaticism while I was there. Were it not for their hospitality, support, and philosophical insights, this paper would not have been written.

[13] Bradley, Seamus and Katie Steele 2015. 'Can free evidence be bad? Value of information for the imprecise probabilist', *Philosophy of Science* 83: 1-28.

[14] Briggs, Rachael 2015. 'Costs of abandoning the sure-thing principle', *Canadian Journal of Philosophy* 45, no.5: 827-40.

[15] Buffon, Georges-Louis Leclerc de 1777. 'Essai d'Arithmétique Morale', in *Supplement a l'Histoire Naturelle*, vol. 4 (Paris: L'Imprimerie Royale).

[16] Chalmers, David 2010. 'The singularity: A philosophical analysis', *Journal of Consciousness Studies*, 17, no. 9: 7-65.

[17] Chen, Daniel, Tobias Moskowitz, & Kelly Shue 2016. 'Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires', *The Quarterly Journal of Economics* 131, no. 3: 1181-242.

[18] d'Alembert, Jean le Rond 1761. *Opuscules Mathématiques ou Mémoires sur différens sujets de Géométrie, de Méchanique, d'Optique, d'Astronomie, Tome Second* (Paris: David).

[19] Dyson, Freeman 1981. 'Life in the universe', unpublished Darwin Lecture, Darwin College, Camrbridge, Nov 10, 1981.

[20] Dyson, Freeman 1979. 'Time without end: Physics and biology in an open universe', *Reviews of Modern Physics*, 51, no. 3.

[21] Easwaran, Kenny 2014, 'Decision theory without representation theorems', *Philosopher's Imprint* 14, no. 27:1-30.

[22] Garrett, Brandon, Gregory Mitechell & Nicholas Scurich 2018. 'Comparing categorical and probabilistic fingerprint evidence', *Journal of Forensic Sciences*, 63, no. 6: 1712-7.

[23] GiveWell 2020. 'GiveWell's cost-effectiveness analyses', *GiveWell*, available at: https://www.givewell.org/how-we-work/our-criteria/cost-effectiveness/cost-effectiveness-models. (Accessed April 2020)

[24] Greaves, Hilary 2017. 'Population axiology', *Philosophy Compass*, 12, no. 11: 12442.

[25] Gurmankin, Levy, & Jonathan Baron 2005, 'How bad is a 10% chance of losing a toe?: Judgments of probabilistic conditions by doctors and laypeople', *Memory and Cognition* 33: 1399–406.

[26] Hájek, Alan 2014. 'Unexpected expectations', *Mind*, 123, no. 490: 533-67.

[27] Hájek, Alan n.d. 'Most counterfactuals are false', unpublished manuscript.

[28] Hájek, Alan & Wlodek Rabinowicz n.d. 'Degrees of commensurability and the repugnant conclusion', unpublished manuscript.

[29] Hanoch, Yaniv *et al.* 2019. 'Reaping the benefits and avoiding the risks: Unrealistic optimism in the health domain', *Risk Analysis* 39, no. 4: 792-804.

[30] Huemer, Michael 2008. 'In defence of repugnance', *Mind* 117.468: 899-933.

[31] Huntley, Nathan, Robert Hable, and Matthias Troffaes 2014. 'Decision making', in T. Augustin *et al.* (eds) *Introduction to Imprecise Probabilities* (New York: John Wiley Sons): 190–206.

[32] Isaacs, Yoaav 2016. 'Probabilities cannot be rationally neglected', *Mind* 125, no. 499: 759-62.

[33] Joyce, James 2011. 'A defence of imprecise credences in inference and decision making', *Philosophical Perspectives* 24: 281-323.

[34] Kahneman, Daniel & Amos Tversky 1982. 'Evidential impact of base rates', in D. Kahneman, P. Slovic & A. Tversky (ed.). *Judgment under uncertainty: Heuristics and biases* (Cambridge: Cambridge University Press): 153–60.

[35] McMahan, Jefferson 1981. 'Problems of population theory', *Ethics* 92, no. 1: 96-127.

[36] McClelland, Gary, William Schulze & Don Coursey 1993. 'Insurance for low-probability hazards: A bimodal response to unlikely events', in *Making Decisions About Liability and Insurance* (Dordrecht: Springer): 95-116.

[37] Mogensen, Andreas 2021. 'Maximal cluelessness', *Philosophical Quarterly* 71.1:141-62.

[38] Monton, Bradley 2019. 'How to avoid maximising expected utility', *Philosophers' Imprint*, 19.

[39] Ng, Yew-Kwang 2012. 'A case for happiness, cardinalism, and interpersonal comparability', *The Economic Journal*, 107: 1848-58.

[40] Parfit, Derek 1984. *Reasons and Persons* (Oxford: Oxford University Press).

[41] Pascal, Blaise 1852 (1669). *Pensées* (Paris: Dezobry et E. Magdeleine).

[42] Pomatto, Luciano, Philipp Strack, & Omer Tamuz 2018. 'Stochastic dominance under independent noise', *arXiv e-prints*, arXiv:1807.06927.

[43] Rabin, Matthew 2000. 'Risk aversion and expected-utility theory: A calibration theorem', *Econometrica*, 68, no. 5: 1281-92.

[44] Sandberg, Anders, *Grand Futures*, unpublished manuscript.

[45] Schoenfield, Miriam 2012. 'Chilling out on epistemic rationality', *Philosophical Studies* 158: 197-219.

[46] Schoenfield, Miriam 2014. 'Decision making in the face of parity', *Philosophical Perspectives* 28.

[47] Schwitzgebel, Eric 2017. '1% skepticism', *Noûs* 51.2: 271-90.

[48] Seidenfeld, Teddy 2004. 'A contrast between two decision rules for use with (convex) sets of probabilities: Gamma-maximin versus E-admissibility', *Synthese* 140: 69–88.

[49] Smith, Nicholas J.J. 2014.'Is evaluative compositionality a requirement of rationality?', *Mind* 123.490: 457-502.

[50] Tarsney, Christian. 'Exceeding expectations: Stochastic dominance as a general decision theory', unpublished manuscript.

[51] Temkin, Larry 2014. *Rethinking the Good* (Oxford: Oxford University Press).

[52] Thoma, Johanna 2018. 'Risk aversion and the long run', *Ethics*, 129, no. 2: 230-53.

[53] Tipler, Frank 1986. 'Cosmological limits on computation', *Internation Journal of Theoretical Physics*, 25, no. 6.

[54] Gilovich, Thomas, Amos Tversky & Robert Vallone 1985. 'The hot hand in basketball: On the misperception of random sequences', *Cognitive Psychology*, 17, no. 3: 295-314.

[55] Tversky, Amos & Daniel Kahneman 1974. 'Judgment under uncertainty: Heuristics and biases', *Science*, 185: 1124–31.

[56] Tversky, Amos & Daniel Kahneman 1983. 'Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment', *Psychological Review*, 90: 293–315.

[57] van Fraassen, Bas 1984. ‘Belief and the will,’ *Journal of Philosophy* 81: 235-56.

[58] White, Roger 2010. ‘Evidential symmetry and mushy credence,’ *Oxford Studies in Epistemology* 3: 161-86.

[59] Wilkinson, Hayden 2020. ‘Infinite aggregation: Expanded addition’, *Philosophical Studies*.

[60] Wilkinson, Hayden. ‘Chaos, add infinitum’, unpublished manuscript.

[61] Wright, Rita 2009. *The Ancient Indus: Urbanism, Economy, and Society*, (Cambridge: Cambridge University Press).