Don't Bet the Farm: Decision Theory, Inductive Knowledge, and the St. Petersburg Paradox*

Sebastian Liu
Draft: 05/2022 [04/2021]
Word Count: 9002
sebastian.liu@princeton.edu

Abstract

An eccentric decision theorist offers you a chance to play the St. Petersburg game at a price: a fair coin is flipped until it lands heads, at which point it is flipped no more, and you're awarded a prize of 2^n dollars, where n is the number of flips. What is it rational for you to pay to play this game? According to orthodox decision theory, you're rationally required to maximize expected value, and since the expected value of the St. Petersburg game is infinite, you're rationally required to pay an arbitrarily large amount for a chance to play. But it doesn't seem as if the game is actually worth very much – it doesn't seem irrational for you to refuse to pay a large finite amount for a chance to play this game. This is the St. Petersburg paradox. A promising suggestion for resolving this paradox maintains that it's rationally permissible for you to ignore sufficiently low probability outcomes in your decision theoretic reasoning. Unfortunately, this suggestion faces a number of serious challenges. The aim of this paper is offer a solution to the St. Petersburg paradox by giving a different account of when it's permissible to ignore certain outcomes. On the present proposal, it's rational for you to ignore outcomes incompatible with your knowledge. When combined with an anti-skeptical commitment to inductive knowledge about fair coins, this proposal resolves the St. Petersburg paradox, and when suitably generalized, also resolves a range of other puzzles generated by high-value, low-probability outcomes.

An eccentric decision theorist offers you a chance to play the St. Petersburg game at a price: a fair coin is flipped until it lands heads, at which point it is flipped no

*I'd like to thank Calvin Baker, Tito Bastianello, Pietro Cibinel, Felipe Doria, Guus Eelink, Adam Elga, Gideon Rosen, Matt Hewson, Ben Holguín, Tom Kelly, Jeremy Khoo, Bradley Monton, Richard Roth, Bernhard Salow, Miriam Schoenfield, Dan Waxman, and Tim Williamson for helpful feedback and discussion about the material here. Thanks also to audience members at Oxford's 25th Annual Graduate Philosophy Conference for questions, and in particular to Teru Thomas for responding to my talk. I'm especially grateful to Lara Buchak and Harvey Lederman for extensive comments on several drafts.

more, and you're awarded a prize of 2^n dollars, where n is the number of flips. What is it rational for you to pay to play this game?

It would certainly be rational for you to pay 1 dollar to play. After all, in the worst case scenario, the coin lands heads on the first flip, in which case you'll be awarded 2 dollars. That's a guaranteed profit of 1 dollar. Surely it would also be rational for you to pay 2 dollars. After all, there's no chance you'll lose any money, for even if the coin lands heads on the first flip, you'll break even. And if you're lucky, the coin will land tails on the first flip, so you'll walk away with a minimum of 2 dollars profit. Perhaps it would be rational for you to pay 3 dollars. After all, you're guaranteed to win 2 dollars, so you can't lose more than 1 dollar. You're then essentially paying 1 dollar for a $\frac{1}{2}$ chance to win at least an additional 2 dollars. And your expected profit given these odds is positive. Would it be rational for you to pay 4 dollars? 10 dollars? 100 dollars?

According to orthodox decision theory, rationality requires you to act in such a way as to maximize expected value. The expected value of a game is calculated by summing the products of the probability of each possible outcome of the game with the value of that outcome. Where T denotes tails and H denotes heads, the mutually exclusive and jointly exhaustive possible outcomes of the St. Petersburg game are the sequences of flips H, TH, TTH, and so on. Let f_i denote the outcome in which the coin is flipped exactly i times (that is, f_i is the sequence of $i-1$ consecutive tails followed by heads). The probability of f_i obtaining is $\frac{1}{2^i}$, since there's a $\frac{1}{2}$ chance that the coin lands heads on the first flip, a $\frac{1}{4}$ chance that the coin lands tails on the first flip and heads on the second flip, a $\frac{1}{8}$ chance that the coin lands tails on the first two flips and heads on the third flip, and so on. The prize, given that the actual sequence of flips is f_i, is 2^i. The expected value of the game, according to orthodox decision theory, is then $\sum_{i=1}^{\infty} \frac{1}{2^i} \times 2^i = \frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \frac{1}{8} \times 8 + \ldots = 1 + 1 + 1 + \ldots$ which is infinite. But if the expected value of the game is infinite, and you should maximize expected value, then rationality requires you to pay any amount to play.

If you're like me, however, then you won't be willing to pay an arbitrarily large amount for this game. Indeed, if you're like me, you'll think that the value of the St. Petersburg game isn't even particularly high. And this doesn't appear to be because there's a failure of rationality or anything. The game simply is *not* worth very much. Or so it seems. But then what's wrong with the standard expected value reasoning? This is the St. Petersburg paradox.[1]

[1] As far as I know, a version of the paradox was first articulated in a letter by Nicolaus Bernoulli to Pierre Rémond de Montmort in 1713, and a refined version of it was subsequently published by Daniel Bernoulli in 1738. See Bernoulli (1738 [1954, p.31]).

1 Ignoring Low Probability Outcomes

One reaction you might have to the St. Petersburg paradox is to maintain that nothing has gone wrong in the standard decision theoretic reasoning and that it really *is* rational to pay any amount to play.[2] This appears difficult to accept. One peculiar feature of the St. Petersburg game is that the game's expected value is always higher than its actual value – however many times the coin is flipped, you'll only be awarded a finite prize, but the expected value of the game isn't finite. Of course, by itself, this is hardly a decisive argument, but the starting point of this paper is that rationality shouldn't require you to pay an infinite amount for this game.

Another reaction you might have to the St. Petersburg paradox is to reject that the value of the game is infinite on the basis that money presumably has diminishing marginal utility.[3] If money has diminishing marginal utility – if the addition of more money to an outcome has less utility the more money that outcome already contains – then the utility of the game can be finite even if the monetary value of it isn't, for given an appropriately chosen function d which models the diminishing marginal utility of money, $\sum_{i=1}^{\infty} \frac{1}{2^i} \times d(2^i)$ converges.[4] This seems right, as far as it goes, but nothing essential depends on the prize being money. Analogous games can be constructed in which the prize is an increase in the number of happy days lived, or raw units of utility, or whatever else *doesn't* have diminishing marginal utility. Still, in these instances, insofar as you think that it isn't rationally required of you to wager an infinite (or large but finite) amount of money to play the St. Petersburg game, you should also think that it isn't rationally required of you to wager an infinite (or large but finite) number of happy days or raw units of utility to play.

A more promising suggestion involves denying that the value of the game is infinite by defending the idea that in your decision theoretic calculations, it's permissible to ignore certain outcomes when those outcomes have sufficiently low probabilities:

> … we might think that we ought not consider the chance of a good outcome in our calculations if that good outcome is extremely unlikely to

[2]For example, Nover and Hájek (2004, p.241) write: "… there is something to be said for the bullet-biting response: 'the game *should* be valued infinitely, and any intuition to the contrary should be dismissed as an artifact of our finite minds not fully appreciating the true nature of the game; we should learn to live with decision theory's verdict'." My proposal will ultimately (intentionally) leave open the possibility that it's rational to pay an infinite amount for the game. My proposal does, however, suggest that paying an infinite amount for the game isn't *uniquely* rational.

[3]This idea can be traced back to Bernoulli (1738 [1954, pp.24-25]), who suggests: "… the determination of the value of an item must not be based on its *price*, but rather on the *utility* it yields… it is highly probable that *any increase in wealth, no matter how insignificant, will always result in an increase in utility which is inversely proportionate to the quantity of goods already possessed*."

[4]For example, given a logarithmic utility function, the expected value of the St. Petersburg game will be evaluated to be finite, as $\sum_{i=1}^{\infty} \frac{1}{2^i} \times \log_2 2^i = \sum_{i=1}^{\infty} \frac{i}{2^i} = 2$.

happen. (Buchak 2013, p.74)

> There is… some threshold such that the agent would not be *irrational* if she simply ignored outcomes whose probabilities lie below that threshold. Hence decision theory, qua theory of ideally rational decision making, must not mandate that she factor in outcomes of *arbitrarily low* probability: that is, that she consider smaller and smaller probabilities *ad infinitum*. (Smith 2014, p.474)

> … for decision-making, small probabilities should be discounted down to zero before maximizing expected utility… (Monton 2019, p.5)

Roughly, the picture is that you can rationally ignore outcomes whose probabilities are low enough (in some specified sense), by treating these outcomes as if they had probability zero.[5] Call this strategy:

> (PROBABILITY-NEGLECT) In your decision theoretic calculations, it's rationally permissible to ignore an outcome O by discounting the probability of O to zero just in case its probability of obtaining is below some threshold.[6]

Further questions about what exactly this threshold is, whether it's objective or subjective, and whether it's sensitive to context or stakes remain unanswered by PROBABILITY-NEGLECT in its current formulation. But to illustrate how a crude version of PROBABILITY-NEGLECT resolves the St. Petersburg paradox, suppose that in evaluating the game, you can rationally ignore any outcome whose probability is less than $\frac{1}{50000}$. Then, since $\frac{1}{2^{16}} < \frac{1}{50000}$, you can rationally ignore the outcomes in which the coin is flipped 16 times or more, treating these outcomes as if they had probability zero. The value of the game given these constraints would be $\sum_{i=1}^{15} \frac{1}{2^i} \times 2^i$, a much more reasonable 15 dollars.

Despite its promise to solve the St. Petersburg paradox, PROBABILITY-NEGLECT faces serious challenges.[7] The most salient one concerns the arbitrariness of the

[5] Early advocates of such a strategy dating back to Bernoulli – see Spiess (1975) – include d'Alembert (1761), Buffon (1777), and Condorcet (1785).

[6] Empirical research suggests that people have difficulty properly accounting for outcomes which have extremely low probabilities. For example, Kahneman and Tversky (1979, pp.282-283) write: "… the simplification of prospects in the editing phase can lead the individual to discard events of extremely low probability… Because people are limited in their ability to comprehend and evaluate extreme probabilities, highly unlikely events are either ignored or overweighted…" A noteworthy consequence if it's permissible to ignore sufficiently low probability is that such behavior, typically assumed to be irrational, could in fact be rational.

[7] The problems I discuss here are merely cursory, and are by no means the only ones faced by PROBABILITY-NEGLECT. See, for example, Parfit (1981), Isaacs (2016), Wilkinson (2022), Beckstead and Thomas (manuscript), and Cibinel (manuscript) for other objections.

threshold. It seems obvious to wonder, for any given threshold, why the threshold is *that* particular value, and not one which is higher or lower.[8]

It's difficult to see how to motivate an objective threshold, one that applies to everyone, for every decision problem, for it's unclear what kinds of facts or considerations would be relevant in determining that threshold value.[9] Any answer seems to at best only replace the question of why some particular threshold is appropriate with the question of why some particular set of facts or considerations are the ones which bear on determining the threshold. Furthermore, for any fixed objective threshold, it's possible to construct a fair lottery in which the probability of any particular ticket winning is below that threshold value. On a straightforward understanding of an objective threshold, you would then be permitted to ignore every outcome. Surely that's absurd, provided that you know that *some* ticket in the lottery will win.

It's also difficult to see how to motivate a subjective threshold, one that is relativized to individuals, and possibly sensitive to specific details of the decision problems, for the worry of arbitrariness is not satisfactorily answered if you're simply permitted to choose whatever threshold you see fit.[10] For one, there clearly needs to be some non-subjective constraint. It would be absurd if you could ignore an outcome that was more likely than not to obtain. For another, the apparent arbitrariness of a subjective standard seems to be in tension with decision theory insofar as it is meant to be normative. If decision theory is in part supposed to prescribe how you *should* act, and you're permitted to decide the threshold, you seem to be able to guarantee the rationality of some actions merely by picking an appropriate threshold value.

None of this is meant to suggest that there aren't reasonable responses or more sophisticated accounts available for a committed defender of PROBABILITY-NEGLECT. Perhaps some version of this proposal can escape these (and other related) objections. However, PROBABILITY-NEGLECT is only a particular implementation of a broader idea:

> (NEGLECT) In your decision theoretic calculations, it can be rationally permissible to ignore certain outcomes.

By itself, NEGLECT is silent on which outcomes you are permitted to ignore. The answer supplied by PROBABILITY-NEGLECT is: those outcomes with sufficiently low

8 This worry is expressed by, for example, Arrow (1951, p.414): "The probability that a head will not appear until the n^{th} toss becomes very small for n sufficiently large; if the occurrence of that event is regarded as impossible for all n beyond a certain value, then the mathematical expectation of return becomes finite, and the [St. Petersburg] paradox is resolved… [PROBABILITY-NEGLECT] seems extremely arbitrary in its specification of a particular critical probability…"

9 It's been suggested by, for instance, Buffon (1777) and Condorcet (1785), that the objective threshold should be determined by considerations related to the probability of dying.

10 Monton (2019), for example, maintains that the threshold must be subjective.

probabilities. But there are other possible answers. Indeed, NEGLECT can be motivated in a more principled way.

2 When are You Permitted to Ignore Outcomes?

To apply orthodox decision theory in your practical reasoning, you must associate, with each possible outcome of your decision problem, some probability. This exercise is straightforward enough in the St. Petersburg game, since it's only natural to assign the objective probability of some sequence of flips as the probability of that outcome for evaluating the expected value of the game. But it becomes virtually impossible to assign objective probabilities sensibly when the outcomes are more complicated and the objective probabilities aren't known. Consequently, the relevant notion of probability for decision theory is typically taken to be subjective probability, or credence (perhaps constrained in certain ways when the objective probabilities are known, as in the St. Petersburg game).[11] By focusing on the partial attitude of credence, orthodox decision theory leaves out the familiar full attitudes of belief and knowledge.[12] Of particular interest for present purposes is knowledge: if knowledge is immaterial to the calculation of expected value, then if (decision-theoretic) rationality is a matter of acting as to maximize expected value, knowledge is immaterial to rationality.[13] But knowledge and rationality appear to be intimately connected:

> If you know that *p*, then it shouldn't be a problem to act as if *p*. If it is a problem to act as if *p*, you can explain why by saying that you don't know that *p*. Suppose you are faced with some decision – do *A* or do *B* – where which of these is better depends on whether *p*. You know that if *p*, *A* is the thing to do, but that if not-*p*, *B* is. To say in one breath, "I know that *p*" and in the next breath, "But I'd better do *B* anyway, even

[11] Sometimes, the relevant probabilities are taken to be your evidential probabilities.

[12] That full attitudes do not factor in standard decision theory is old news. Some take this to suggest that, at least for the purposes of decision making, full attitudes like belief are completely irrelevant. For example, Kaplan (1996, p.100) writes: "… if the best that can be said about belief is that it is not a state of confidence, if the best that can be said is that it has nothing to do with decision making or the inquiry that serves your decision making when you harbor the values with which we have imagined you are imbued, then it may be because belief *is not anything at all*." For discussion and dissent, see among others Harsanyi (1985), Bratman (1987), Lance (1995), Smithies (2012), Wedgwood (2012), Ross and Schroeder (2014), and Staffel (2019).

[13] A related point is made by Hawthorne and Stanley (2008, p.590): "Our ordinary conceptual scheme suggests a connection between reasons and knowledge that is altogether ignored by the standard decision theory…" Of course, it might be maintained that knowledge is tightly linked to credence 1, in which case knowledge does play some role in decision theory, if only, in a sense, secondarily. I won't take a stance here on whether knowledge entails maximal credence, though see below for further discussion.

> though I know that A is the thing to do if p" seems incoherent. (Fantl and McGrath 2002, p.72)
>
> A standard use of knowledge attributions is to justify action… to say that an action is based on knowledge is to declare that the action has met the expected norm. (Stanley 2005, p.10)
>
> … it is legitimate to write something onto a decision table iff the decision maker knows it to be true… it is legitimate to leave a possible state of the world off a decision table iff the decision maker knows it not to obtain. (Weatherson 2012, p.77)

It seems clear enough that knowledge bears on rationality in some way. Less clear, however, is how exactly knowledge bears on rationality.[14] But if something in this vicinity is correct, it's tempting to think that knowledge (at the very least) constrains the set of outcomes you're required to consider in your decision theoretic calculations – that is, for the purposes of decision making, you're only required to consider the outcomes which are compatible with your knowledge. This motivates a different implementation of NEGLECT. In particular:

> (KNOWLEDGE-NEGLECT) In your decision theoretic calculations, it's rationally permissible to ignore an outcome O by excluding O from the set of possible outcomes whenever O is incompatible with your knowledge.[15]

Knowledge-norms figure prominently across a wide range of debates including assertion, belief, evidence, disagreement, and legal proof.[16] KNOWLEDGE-NEGLECT fits well with this broad knowledge-centric package of views. The norm is relatively modest, for it merely states a sufficient condition for when you're permitted to ignore certain outcomes. In some sense, KNOWLEDGE-NEGLECT might seem almost trivial: *of course* you can permissibly ignore outcomes which you *know* won't obtain when making decision theoretic calculations. Still, some remarks will be helpful in bringing out the intuitive force of this norm.

According to views on which knowledge entails credence 1, KNOWLEDGE-NEGLECT immediately follows. This is because if you know that O won't obtain, your credence that O will obtain will be 0 (assuming your credences obey the standard

[14] For further discussion, see for instance Williamson (2005), Hawthorne and Stanley (2008), Moss (2013), Weisberg (2013), Ross and Schroeder (2014), and Schulz (2015).

[15] Other possible implementations of NEGLECT include ignoring outcomes which for all you know, you know won't obtain, or which you outright believe won't obtain. I won't discuss these possibilities here. Knowledge might bear on constraining outcomes in other ways as well. For instance, Levi (1980, pp.3-5) writes: "X's knowledge at t serves as a standard for distinguishing truth-value-bearing hypotheses whose truth is a serious possibility according to X at t from those whose truth is not a serious possibility according to X at t… h is a serious possibility according to X at t if and only if h is consistent with his corpus of knowledge at t."

[16] See, for example, Williamson (2000) on assertion, belief, and evidence, Hawthorne and Srinivasan (2013) on disagreement, and Moss (2021) on legal proof.

probability axioms) and so O will be effectively ignored. This welcome result is limited by its reliance on the vexed relationship between knowledge and credence.

Fortunately, the idea that you can ignore outcomes which are incompatible with your knowledge in your decision theoretic calculations enjoys significant appeal independently of whether or not knowledge entails maximal credence.[17] On the standard way of modeling knowledge, you know a proposition just when that proposition is true in every epistemically accessible world. So if you know that an outcome O won't obtain, in no epistemically accessible world will O obtain, that is, in every epistemically accessible world, O will fail to obtain.[18] But then why should you be required to consider O in your decision theoretic reasoning?

This point can be further illustrated by an example. Suppose that I am deciding whether to serve you a salad or a peanut butter sandwich for lunch, and suppose that I am very confident that you enjoy sandwiches over salads. Consider a case in which I know you're not allergic to peanuts. Here, it's completely natural for me to serve you the peanut butter sandwich. If asked why I decided to serve you the sandwich instead of the salad, it's perfectly reasonable for me to respond by citing my confidence in your preference of sandwiches over salads, and if asked why I didn't consider outcomes in which you're allergic to peanuts, it's perfectly natural for me to cite my knowledge that you're not allergic to peanuts. A response to the effect that it's at least logically or metaphysically possible that you've recently developed a peanut allergy, or that I've badly misremembered and therefore that I should have non-zero credence that you're allergic, is uncompelling as criticism that I should have factored in these outcomes in my deliberation.[19] If I in fact know that you're not allergic, it seems that these kinds of possibilities become irrelevant. This is in stark contrast with a case in which I lack knowledge about whether you're allergic to peanuts. Here, even if I'm highly confident that you're not allergic, I would be blameworthy for neglect if I served you the peanut butter sandwich. My high confidence is not sufficient for ignoring the outcomes in which you are allergic to peanuts, and the inviting explanation is because I lack the relevant knowledge.[20]

[17] Interestingly, empirical evidence suggests that knowledge figures prominently in everyday evaluations of behavior. For example, Turri (2017, p.2253): "Alicia is at Metro Beach with her children. She examines the water and concludes that it is safe for swimming. Should Alicia allow her children to go swimming? Researchers found that people's response to this question depends on whether Alicia 'knows,' 'thinks,' or 'is certain' of her conclusion. When she 'thinks' or 'is certain,' people disagreed that she should allow her children to go swimming. But when she 'knows,' people tended to agree..." If KNOWLEDGE-NEGLECT is correct, the rationality of such reliance on knowledge is vindicated.

[18] The factivity of knowledge ensures that the actual world is always epistemically accessible, so if you know that O won't obtain, then in fact, O won't obtain.

[19] If knowledge entails credence 1, then given my knowledge, my credence that you're allergic is (or should be) 0.

[20] An interesting case concerns whether it's rational for you to buy (reasonably priced) insurance. On the one hand I can sometimes get into the frame of mind where I'm happy to agree that you're typically in a position to know that you won't get seriously ill within the next year. On the other

Orthodox decision theory requires you to assign probabilities to all of the mutually exclusive and jointly exhaustive possible outcomes in your decision problem. It's silent about how these possible outcomes are related to your knowledge. Expected value is then calculated by summing the products of these probabilities with the values of their corresponding outcomes. KNOWLEDGE-NEGLECT maintains that the only outcomes required for the calculation of expected value are the ones *consistent with your knowledge*.[21]

In comparison with PROBABILITY-NEGLECT, KNOWLEDGE-NEGLECT appears to be a better motivated implementation of NEGLECT, and the norm suggests a very natural way in which knowledge interacts with (decision-theoretic) rationality. Unlike PROBABILITY-NEGLECT, KNOWLEDGE-NEGLECT is sensitive to the way in which the relevant probabilities are generated. The norm does not allow you to ignore the outcomes in which you've lost a fair lottery, provided that lottery propositions are unknowable, but it does permit you to ignore outcomes involving you being a brain-in-a-vat, provided that you know that you're not a brain-in-a-vat. Therefore, unlike PROBABILITY-NEGLECT, KNOWLEDGE-NEGLECT does not solely depend on the probabilities of the outcomes. Whether you can ignore an outcome depends only on whether you know that it won't obtain. Because knowledge is factive, the actual outcome can't be ignored, regardless of how low its probability is. A consequence is that, unlike PROBABILITY-NEGLECT, KNOWLEDGE-NEGLECT does not by itself offer a solution to the St. Petersburg paradox, for by itself, it is silent on which outcomes (if any) can be ignored in the St. Petersburg game.

3 Knowledge and Coin Flips

Suppose that you're given a coin which you know with equal probability is either fair or double-tailed. Provided that the coin is in fact double-tailed, it seems that, if knowledge from induction is at all possible, you can determine that it is not fair by flipping it 1000 times (for example) and observing it land tails every time:

> …if you could ever learn anything non-trivial about objective chances,

hand, following, for example, Hawthorne and Stanley (2008), I'm more often in the frame of mind where I'm inclined to deny that you have such knowledge.

[21] It may be helpful to examine formal implementations of KNOWLEDGE-NEGLECT. Given a decision problem, let Ω be the set of possible outcomes, and where $O \in \Omega$, let $P(O)$ be the probability of O, and let $u(O)$ be the value of O. Orthodox decision theory calculates expected value as:

$$\sum_{O \in \Omega} P(O) \times u(O)$$

One way of thinking about KNOWLEDGE-NEGLECT is understanding it to manipulate Ω by replacing Ω with Ω_K, the set of possible outcomes compatible with your knowledge. Another way of thinking about KNOWLEDGE-NEGLECT is understanding it as permitting you to conditionalize (in the standard way) on your knowledge.

> you could learn that a certain double-headed [double-tailed] coin is not fair by flipping it repeatedly, seeing it land heads [tails] each time, and eventually inferring that it is not fair. (Dorr, Goodman, and Hawthorne 2014, p.283)

To be sure, it's *possible* for a fair coin to land tails 1000 times in a row. The probability of that sequence (like any other sequence of 1000 flips) is $\frac{1}{2^{1000}}$. When, by a stroke of extreme luck, you have a fair coin which does land tails every time when it is flipped 1000 times, you wouldn't know that it's not fair, as knowledge is factive. But it's not obvious that the mere possibility that a fair coin can produce a sequence of 1000 tails is good reason for denying that when your coin is in fact double-tailed, you can have knowledge that your coin isn't fair.[22] Your knowledge is often compatible with some skeptical hypotheses when these hypotheses are true and you're in the 'bad' case, though this is hardly evidence for thinking that your knowledge is still compatible with these skeptical hypotheses when these hypotheses are false and you're in the 'good' case. Maintaining that you could *never* know that a double-tailed coin is not fair by observing the outcomes of its flips seems tantamount to accepting a radical form of skepticism about inductive knowledge:

> [one] reason to think that we can directly rule out some sequences of coins in advance is that it would otherwise be very hard to account for knowledge acquired by induction. (Bacon 2014, p.377)
>
> …if we deny [that you know that not all the flips of the fair coins landed tails] then we would most probably need to discount any of our knowledge that has a probabilistic evidential basis, which results in a wide-ranging skepticism. (Rothschild and Spectre 2018, p.473)

Suppose you are ignorant that you have a double-tailed coin. Then, if you can know that it's not fair *after* observing the outcomes of 1000 flips, plausibly *before* flipping the coin, you can know the conditional: if the coin is fair, it won't land tails 1000 times in a row. It would be striking if, before flipping your coin, it's consistent with your knowledge that a fair coin will land tails 1000 times consecutively, but once you've flipped your coin, your knowledge becomes incompatible with a fair

[22]The judgment that you can have this kind of inductive knowledge isn't unique to coin flips. Here are two more cases. Suppose you're wondering whether I'm a card cheat. After playing 100 rounds of poker with me, you notice that I've had a royal flush every time. This is of course possible – the probability of being (fairly) dealt a royal flush 100 times in a row is $\frac{1}{649740^{100}}$. But it seems that, when I am cheating, you can come to know that I am on the basis of this extremely low probability. Or suppose you suspect that I've been rigging the local one-million ticket lottery. It seems you can confirm your suspicions and come to know that I am not merely getting very lucky every time when you learn that I have won the past 100 lotteries in a row. This is despite the fact that the probability of winning 100 fair one-million ticket lotteries consecutively is $\frac{1}{1000000^{100}}$. In these examples, it's exceedingly tempting to grant that you do have the relevant knowledge when I am cheating at poker and have been rigging the lottery.

coin landing tails 1000 times consecutively.[23] So setting aside inductive knowledge skepticism, you have:

> (FAIR-COIN-KNOWLEDGE) You know that a fair coin will not land tails 1000 times in a row when flipped repeatedly, provided that it doesn't.[24]

If you're a not a skeptic of the relevant variety, you should accept that inductive knowledge is possible.[25] And if you accept that inductive knowledge is possible, you should think that you can come to know, of a double-tailed coin, that it's not fair upon flipping it 1000 times and seeing it come up tails every time. And if you think that such knowledge about coins is possible after seeing the outcomes of the flips, you should accept that you can have FAIR-COIN-KNOWLEDGE.

4 Solving the St. Petersburg Paradox?

Put succinctly, the thesis of this paper is that *given* both KNOWLEDGE-NEGLECT and FAIR-COIN-KNOWLEDGE – both of which are individually compelling – the St. Petersburg paradox can be resolved. Here is why.

Suppose that in your decision theoretic calculations, you can permissibly ignore outcomes which you know won't obtain. And suppose that you can know of a fair

[23] Perhaps there's room here for resistance. For example, according to certain dogmatist views about perception, you receive some (defeasible) justification to believe p only when you have an experience as if p. Similarly, it might be maintained that you can come to know that a coin isn't fair only when you see it land tails 1000 times in a row. See especially Pryor (2000), Wright (2004), and White (2006) for discussion. Let me sketch what I take to be one highly undesirable feature of this position: before flipping, you would have to think to yourself 'for all I know, a fair coin will land tails every time when flipped 1000 times in a row', but after seeing the coin land tails 1000 times in a row, you would have to think to yourself 'now it is inconsistent with my knowledge that the coin is fair'. What would afford you this knowledge if you couldn't know the conditional prior to flipping the coin?

[24] This is perhaps slightly too quick, for if knowledge requires safe belief, and the fair coin will land tails 999 times in a row, then you arguably wouldn't know that it won't land tails 1000 times in a row. A qualification of safety can be added, though it is largely immaterial to the discussion, for it would be a mistake to place any significance of the number of flips being exactly 1000. It's sufficient for my purposes that there's some finite sequence of length n of all tails such that you know that a fair coin won't land tails n times in a row. Moreover, nothing depends on questions about whether n is vague. It may be that for some finite sequences of all tails, it's vague whether you can know that the coin is not fair upon seeing the outcomes of flips match those sequences. What matters is that there's some large enough n beyond this possibly vague interval such that it is definite that you have the relevant knowledge.

[25] I don't intend for the discussion to depend on any particular characterization of inductive knowledge. A helpful way of understanding it is offered by Goodman and Salow (manuscript, pp.1-2): "In theorizing about what people know, it is often productive to factor their knowledge into two components: knowledge of some starting points (such as instrument readings, memories, perceptual appearances, and other background certainties) and *inductive* knowledge that goes beyond these starting points. Call these starting points a person's *evidence*. Inductive knowledge is then knowledge that goes beyond one's evidence."

coin, that when flipped 1000 times, it won't land tails every time, provided that it in fact doesn't. Then when faced with the St. Petersburg game, when you have FAIR-COIN-KNOWLEDGE, by KNOWLEDGE-NEGLECT, you're rationally permitted to ignore all of the outcomes in which the coin is flipped at least 1000 times. The value of the game in these instances therefore has an upper-bound of 1000 dollars.[26]

But what of the cases, however improbable, in which the fair coin *does* land tails at least 1000 times in a row? In these unlikely scenarios, by the factivity of knowledge, you wouldn't have FAIR-COIN-KNOWLEDGE, and so you couldn't permissibly ignore the outcome in which the coin is flipped 1000 times.

Importantly, in these instances, it's consistent with KNOWLEDGE-NEGLECT that the expected value of the game is greater than 1000 dollars. But the paradoxical result that the game is worth an infinite amount of money can nevertheless be resisted, for even if the coin is going to be flipped more than 1000 times, you can still presumably know that it won't land tails 5000 (or 10000, or 1000000, or...) times in a row (provided that it doesn't). So despite the fact that when extremely improbable outcomes do obtain, they cannot be permissibly ignored, there will be other, even more improbable outcomes incompatible with your knowledge which can be permissibly ignored.[27]

The intuitive verdict that it can be rational not to pay an arbitrarily large amount for the St. Petersburg game is vindicated. For every finite sequence of flips, there will only be a finite number of outcomes you're required to consider in your decision theoretic calculations.[28] There's therefore a motivated reason to maintain that the value of the St. Petersburg game is finite. Paradox resolved.[29]

[26]Isn't the game intuitively worth a lot less than 1000 dollars? Sure. 1000 flips seems to me to be an exceedingly safe upper-bound for when you can come to know that the coin is not fair. If pressed, I'd say you can come to know after around 30 flips. If so, the St. Petersburg game will at most be worth 30 dollars (when you have the relevant piece of knowledge). Moreover, although certain other decision theories like risk-weighted expected utility theory can't, by themselves, resolve the St. Petersburg paradox, when combined with the present proposal, the game might be evaluated to be worth much less. See especially Buchak (2013).

[27]A noteworthy consequence of present proposal is that the expected value depends in part on what the actual sequence of flips is. Depending on how the coin in fact lands, the expected value of the game may be higher or lower. This should be expected, if knowledge is to play some role in constraining rationality, as knowledge is factive.

[28]The interesting case is the one in which the fair coin never lands heads. See especially Williamson (2007). The sequence of all tails is surely *possible* – though it's unclear what the prize here would be, given the standard description of the St. Petersburg game. However, the sequence of all tails won't be a problem for the present proposal if it's simply stipulated that the prize given this sequence is finite.

[29]What about a modified version of the St. Petersburg paradox? Suppose the St. Petersburg game is played with a random sequence of heads and tails: a fair coin will be flipped until it diverges from this random sequence, at which point the coin will be flipped no more and you'll be awarded a prize of 2^n dollars, where n is the number of flips. The thought is that, even if you can have FAIR-COIN-KNOWLEDGE, this doesn't extend to every sequence of heads and tails, so you should be willing to pay an arbitrarily large amount for this variation of the game. I won't be able to address this worry in

Or is it? You may feel uneasy about this solution for several reasons.

Here is one worry you might have: if the question of whether some outcome can be permissibly ignored depends on whether that outcome is compatible with your knowledge, and your knowledge depends in part on whether that outcome in fact will obtain, then it seems that the resulting decision theory will be less than fully action-guiding, for you won't always be in a position to know whether some outcome can be permissibly ignored. When you have a fair coin that will land tails 1000 times in a row, things won't seem any different to you, from your perspective, than when you have a fair coin that will land heads after (for instance) the third flip, and since your knowledge in these two cases differ, the outcomes you can permissibly ignore will differ, so the game's expected value will differ, and therefore what it's rational for you to pay for a chance to play the game will differ.

If a genuine desideratum for decision theory is that it be action-guiding in *every* circumstance, then orthodox decision theory should not be thought to have an obvious advantage over one in which you're permitted to ignore outcomes you know won't obtain, for it's unclear whether orthodox decision theory itself is action-guiding in this sense. Decision problems which are highly sensitive to your credences prove to be problematic unless you're always in a position to access your credences. Given anti-luminosity considerations, such an assumption about the accessibility of credences looks dubious.[30] If, as is likely, your credences are non-luminous, orthodox decision theory too, will fail to be strictly action-guiding.

Action-guidance can be conceived more leniently – perhaps what is required is merely that decision theory be *typically* action-guiding. If the (positive) knowledge iteration principle holds, then whenever it is permissible for you to ignore an outcome, you'll be in a position to know that it's permissible for you to ignore that outcome, and so it's plausible that KNOWLEDGE-NEGLECT satisfies this more lenient requirement of action-guidance.[31] But even if knowledge doesn't iterate, generally, you still will know enough about what you do (and do not) know to determine which outcomes you can (and cannot) permissibly ignore.

A different worry you might have concerns ignoring outcomes incompatible with your knowledge in high-stakes cases: if it's permissible for you to ignore an outcome given that you know that this outcome won't obtain, then it's permissible for you to ignore it regardless of what is at stake. But consider the following bet: if

---

detail, but here's one consideration in favor of thinking that FAIR-COIN-KNOWLEDGE extends to these cases. Suppose that you're given a coin which you know with equal probability is either fair, or rigged as to result in a particular sequence of heads and tails when flipped. Provided that the coin is in fact rigged, it seems that you can come to know that it is not fair by flipping it 1000 times and observing it match the particular sequence.

[30]For anti-luminosity arguments in a probabilistic setting, see especially Williamson (2008).

[31]However, KNOWLEDGE-NEGLECT still won't be perfectly action guiding, since the negative introspection principle – that ignorance entails knowledge of ignorance – is surely false. See in particular Stalnaker (2006). Ordinarily, when a fair coin will in fact land tails 1000 times in a row, you won't be in a position to know that you lack FAIR-COIN-KNOWLEDGE.

a fair coin, to be flipped 1000 times, doesn't land tails every time, you win 1 dollar, but otherwise, the world will be destroyed. If you have FAIR-COIN-KNOWLEDGE, then given KNOWLEDGE-NEGLECT, it's rationally permissible for you to take this bet, since you're permitted to ignore the outcome of all tails. But isn't this wrong? It seems that if the destruction of the world is at risk, it would not be rational for you to take this bet. When you can permissibly ignore an outcome appears to be influenced by what is at stake.

It's unclear whether this sort of problem is specific to KNOWLEDGE-NEGLECT. Consider the following bet: if a fair coin, to be flipped an infinite number of times, lands tails only a finite number of times, you win 1 dollar, but otherwise, the world will be destroyed. According to orthodox decision theory, you're rationally required to accept this bet. But isn't this wrong? It seems that if the destruction of the world is at risk (it's certainly a possibility that the coin will land tails only a finite number of times), it would not be rational for you to take this bet.

Perhaps in these kinds of cases, it is rational to take the bet. When a fair coin is flipped an infinite number of times, outcomes in which the coin lands tails only a finite number of times, after all, have probability 0, though of course these outcomes are still metaphysically possible. But if this response is available to the defender of orthodox decision theory, why isn't an analogous version of it available to the defender of KNOWLEDGE-NEGLECT? When you have FAIR-COIN-KNOWLEDGE, outcomes in which the coin is flipped 1000 times or more are, after all, incompatible with your knowledge, though of course these outcomes are still metaphysically possible. Perhaps a tempting suggestion is to distinguish between what it is rational for you to do and what you would be criticizable for doing. Since the outcomes in which the coin lands tails only a finite number of times when it is flipped an infinite number of times have probability 0, it's rational for you take the relevant bet. However, since such outcomes are nevertheless possible, you would be criticizable for doing so, given that what is at stake is the destruction of the world. But if this response on behalf of orthodox decision theory is satisfactory, an analogous version of it should also be satisfactory for KNOWLEDGE-NEGLECT: while it is rational for you to ignore the outcome of 1000 consecutive tails when it is incompatible with your knowledge, you would be criticizable for doing so, given that what is at stake is the destruction of the world.[32]

There is also a broader defensive maneuver available to address these and related worries. To the extent that action-guidance and high-stakes are objections to KNOWLEDGE-NEGLECT, they're certainly not unique or specific to it. Any norm appealing to knowledge faces the same set of alleged problems. For instance, the knowledge norm of belief states that you ought to believe all and only what you

[32]Since KNOWLEDGE-NEGLECT doesn't *obligate* you to ignore outcomes incompatible with your knowledge, it is open to a defender of KNOWLEDGE-NEGLECT to maintain that precisely in these circumstances, you would be criticizable (though perfectly rational) for ignoring some outcomes incompatible with your knowledge.

know. Insofar as the knowledge norm correctly describes when it is appropriate to believe something, it would be unreasonable to require that you always be in a position to determine whether your belief is permitted (assuming that knowledge is non-luminous); likewise it would be unreasonable to demand that the permissibility of the belief reflect the stakes involved for the belief (assuming that knowledge is insensitive to stakes). The point is not to insist that the knowledge norms are correct. Rather, the point is to highlight that KNOWLEDGE-NEGLECT is in good company: it is no more objectionable than any of the other knowledge norms.

You might still have other worries about KNOWLEDGE-NEGLECT.[33] And you might also have worries about FAIR-COIN-KNOWLEDGE.[34] The cursory discussion here isn't meant to address all potential objections. But a denier of KNOWLEDGE-NEGLECT or FAIR-COIN-KNOWLEDGE faces a general challenge: if you deny KNOWLEDGE-NEGLECT, you should explain what role (if any) knowledge plays in (decision-theoretic) rationality, and if you deny FAIR-COIN-KNOWLEDGE, you should explain how (if at all) radical skepticism (about inductive knowledge) can be avoided.

Assuming KNOWLEDGE-NEGLECT is right, rationality depends not just on your credences and the values of the possible outcomes of your decision problem, for it also depends on your knowledge. Assuming FAIR-COIN-KNOWLEDGE is right, you can have knowledge that a fair coin won't land tails 1000 times in a row when flipped repeatedly, provided that it doesn't. Given both KNOWLEDGE-NEGLECT and FAIR-COIN-KNOWLEDGE, you're not rationally required to pay an arbitrarily large amount for the St. Petersburg game.[35]

---

[33]See for example Greco (2013), Mueller and Ross (2017), Comesaña (2019), and Fassio and Gao (2021) on problems for knowledge-based decision theories.

[34]Some, for instance, Goodman (1953), are suspicious of the possibility of inductive knowledge. A pressing issue for FAIR-COIN-KNOWLEDGE is that it appears to resemble a lottery proposition. Can this can be resisted? Rothschild and Spectre (2018, p.475) write: "... it would be a mistake to think that theoretical consistency requires us to take the same attitude toward coin propositions as to lottery propositions." I don't want to take a stand on this here, though I'm somewhat hesitant to agree, since it's possible to construct a fair lottery on the basis of coin flips, by mapping each sequence to some unique ticket. If this is right, then although perhaps initially surprising, whether lottery propositions are knowable will depend in part on how the lottery is constructed. See also in particular Bacon (2020).

[35]Especially noteworthy is the fact that this solution to the paradox does not depend essentially on any assumptions about the relationship between knowledge and credence, or credence and objective probability. All that is required is a claim about the role of knowledge in determining when you can permissibly ignore certain outcomes and an anti-skeptical claim about inductive knowledge concerning coin flips. There is an interesting putative tension between FAIR-COINS-KNOWLEDGE, and the following two principles:

(KNOWLEDGE-CREDENCE) $Cr(p \mid Kp) = 1$

(PRINCIPAL-PRINCIPLE) $Cr(p \mid P(p) = x \wedge E) = x$

Roughly, KNOWLEDGE-CREDENCE states that your credence in p, conditional on you knowing that p should be 1, and the PRINCIPAL-PRINCIPLE states that if at time t, your admissible evidence E is compatible with your evidence that the objective chance of p, $P(p)$, is x, then your credence in p should

5 On Two Related Puzzles: The Pasadena and Petrograd Games

The proposal here can be extended to offer straightforward diagnoses of related decision theoretic puzzles, including the Pasadena game and the Petrograd game.

The Pasadena Game. The Pasadena game modifies the St. Petersburg game by awarding you a prize of $(-1)^{n-1} \times \frac{2^n}{n}$ dollars, where n is the number of flips.[36] Consequently, if n is odd, you're paid some amount of money, but if n is even, you're required to pay some amount of money. Orthodox decision theory calculates the expected value of this game as $\sum_{i=1}^{\infty} \frac{1}{2^i} \times (-1)^{i-1} \times \frac{2^i}{i} = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$ When the terms are arranged in this order, the sum converges to $\ln 2$. But interestingly, the sum can be made to converge to any value, or to diverge to positive or negative infinity, when the terms are suitably rearranged. The Pasadena game is supposedly problematic for orthodox decision theory, since it's unclear why one arrangement of the terms should be privileged over another, and so it's unclear whether game has an expected value at all.

The Pasadena game takes advantage of the fact that $\sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i}$ is conditionally convergent.[37] However, given KNOWLEDGE-NEGLECT and FAIR-COIN-KNOWLEDGE, when the coin in fact is not flipped more than 1000 times, the value of the game has an upper-bound of $\sum_{i=1}^{1000} \frac{(-1)^{i-1}}{i}$, which is not conditionally convergent (and is approximately 0.69).

The Petrograd Game. The Petrograd game modifies the St. Petersburg game by awarding you a prize of $2^n + 1$ dollars, where n is the number of flips.[38] Intuitively, this game is worth more than the St. Petersburg game, so if you're offered a chance to play either of these games with the same fair coin, you should choose the Petrograd game over the St. Petersburg game. But orthodox decision theory fails to deliver this verdict, as the expected value of the Petrograd game is not higher than

be x. For if you have FAIR-COINS-KNOWLEDGE, while KNOWLEDGE-CREDENCE requires you to have credence 0 that the coin will land tails every time, the PRINCIPAL-PRINCIPLE requires you to have credence $\frac{1}{2^{1000}}$ that the coin will land tails every time. See Lewis (1980) on the PRINCIPAL-PRINCIPLE, and especially Bacon (2014) on this tension more generally. My arguments here don't depend on taking any particular stand on which (if any) of KNOWLEDGE-CREDENCE or the PRINCIPAL-PRINCIPLE to give up. If your knowledge should receive credence 1, then there's a sense in which orthodox decision theory is correct, though the probabilities for the outcomes in the St. Petersburg game in which the coin is flipped at least 1000 times should be 0, when you have FAIR-COINS-KNOWLEDGE. If you should conform your credences to the known objective chances, then orthodox decision theory wrongly neglects the fact that some outcomes you should have non-zero credence in are nevertheless incompatible with your knowledge.

[36] See Nover and Hájek (2004).

[37] A series a_n is conditionally convergent if $\sum a_n$ converges while $\sum |a_n|$ diverges. By the Riemann series theorem, any infinite series (of reals) which is conditionally convergent can be made to converge to any (real) value, or to diverge, by rearranging the terms.

[38] See Colyvan (2008).

the expected value of the St. Petersburg game, as neither is finite.

Given both KNOWLEDGE-NEGLECT and FAIR-COIN-KNOWLEDGE, it's obvious why you should prefer the Petrograd game over the St. Petersburg game: if you know that the fair coin won't be flipped more than 1000 times, there are only a finite number of outcomes to consider in your decision theoretic calculations, and for any finite number of flips, the expected value of the Petrograd game is higher than that of the St. Petersburg game.[39]

6 High Value Outcomes with Low Probabilities

The St. Petersburg, Pasadena, and Petrograd games are allegedly problematic for orthodox decision theory because they involve certain outcomes with extremely high values and extremely low probabilities. These low probabilities were generated by the flips of a fair coin, but nothing essential seems to depend on that particular feature of the puzzle. The more general feature of these puzzles involves enormous value outcomes with extremely tiny probabilities.

Here's an example of such a puzzle that doesn't involve the flips of a fair coin. In Pascal's Mugging, you're stopped on the streets and asked to hand over your wallet to a stranger claiming to be an Operator of the Seventh Dimension.[40] The stranger promises that, in exchange for your wallet, you'll be awarded handsomely in the near future – the prize will be much, much more valuable than whatever is currently in your wallet. The stranger offers you the following argument: surely your credence that this individual is from the Seventh Dimension and can give you an incredibly valuable prize is greater than zero (don't you only assign credence 0 to necessary falsehoods?); supposing that your credence is ϵ (where $\epsilon > 0$), the stranger assures you that you'll be awarded a prize worth $\frac{10^{50}}{\epsilon}$ times more than the value of the contents in your wallet, so according to orthodox decision theory, maximizing expected value requires you to hand over your wallet.

KNOWLEDGE-NEGLECT offers a potential solution to Pascal's Mugging in a similar way that it offers a potential solution to the St. Petersburg paradox: given KNOWLEDGE-NEGLECT, if your knowledge is incompatible with the possibility that the stranger is from the Seventh Dimension and is able to make good on the dubious promise, then you can ignore that outcome in your decision theoretic calculations. And, on pain of skepticism, it does seem like you can have such knowledge. The evidence you have about the (non)-existence of beings from the Seventh Dimension seems to afford you inductive knowledge that this stranger is not a wealthy and generous inhabitant of the Seventh Dimension. Skepticism aside, this sort of knowledge is possible, regardless of whether knowledge entails maximal credence. And

[39]Moreover, the present solution delivers the verdict that the Petrograd game is worth exactly 1 dollar more than the St. Petersburg game.

[40]See Bostrom (2009). See also Balfour (2021).

if this sort of knowledge is possible, when you have it, given KNOWLEDGE-NEGLECT, you can permissibly ignore the outcome in which you'll be awarded an incredibly valuable prize for giving up your wallet in your decision theoretic reasoning.

The general kind of response to decision theoretic puzzles that involve enormous value outcomes with extremely tiny probabilities sketched here is to combine KNOWLEDGE-NEGLECT with an anti-skeptical principle that relies on a relevant piece of everyday knowledge. In the case of the St. Petersburg paradox and its variants, the anti-skeptical principle concerns what you can know about the outcomes of the flips of a fair coin, and in the case of Pascal's Mugging, the anti-skeptical principle concerns what you can know about individuals claiming that they're from the Seventh Dimension.

However, even if you are committed to both KNOWLEDGE-NEGLECT and the relevant anti-skeptical principle, the scope of the proposal should not be exaggerated or thought to extend to every decision theoretic problem that contains high-value, low-probability outcomes. What is essential for KNOWLEDGE-NEGLECT is your knowledge, and your knowledge is not always incompatible with low-probabilities. To take an example with significant practical import: if it's consistent with your knowledge that some calamity (for instance, malevolent artificial intelligence) is an existential threat, then that outcome cannot be permissibly ignored in your decision theoretic calculations, regardless of how low its probability is.[41] Whether these sorts of cases are problematic for decision theory in the same way that the St. Petersburg paradox is problematic is open.[42] For now, some solace must be found in knowing that there's compelling reason to think rationality does not require you to be swindled by eccentric decision theorists offering exotic coin games and self-professed Seventh Dimension muggers.

References

[1] Arrow, Kenneth J. 1951. Alternative Approaches to the Theory of Choice in Risk-Taking Situations. *Econometrica*, 19, pp. 404-437.

[2] Bacon, Andrew. 2014. Giving Your Knowledge Half a Chance. *Philosophical Studies*, 171, pp. 373-397.

[3] Bacon, Andrew. 2020. Inductive Knowledge. *Noûs*, 54, pp. 354-388.

[4] Balfour, D. 2021. Pascal's Mugger Strikes Again. *Utilitas*, 33, pp. 118-124.

[5] Beckstead, Nick and Teruji Thomas. manuscript. A Paradox for Tiny Probabilities and Enormous Values.

[41] This is particularly relevant for longtermists, who are concerned with how well the far future goes. See especially Greaves and MacAskill (manuscript).

[42] It might be helpful to think about these cases as ones in which the relevant probabilities are generated by conditionalizing on your knowledge. Here, KNOWLEDGE-NEGLECT will be silent, since all of the outcomes will, by construction, be compatible with your knowledge. At least to my mind, however, these kinds of cases are intuitively less troubling.

[6] Bernoulli, Daniel. 1738. Specimen Theoriae Novae de Mensura Sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 5, pp. 175-192. English translation: 1954. Exposition of a New Theory on the Measurement of Risk. *Econometrica*, 22, pp. 23-36.

[7] Bostrom, Nick. 2009. Pascal's Mugging. *Analysis*, 69, pp. 443-445.

[8] Bratman, Michael. 1987. *Intention, Plans and Practical Reason*. Cambridge, MA: Harvard University Press.

[9] Buchak, Lara. 2013. *Risk and Rationality*. Oxford: Oxford University Press.

[10] Buffon, Georges-Louis Leclerc de. 1777. Essai d'Arithmétique Morale. In his *Supplement a I'Histoire Naturelle, Volume IV*. Paris: L'Imprimerie Royale.

[11] Cibinel, Pietro. manuscript. A Dilemma for Nicolausian Discounting.

[12] Colyvan, Mark. 2008. Relative Expectation Theory. *Journal of Philosophy*, 105, pp. 37-44.

[13] Comesaña, Juan. 2019. A Note on Knowledge-First Decision Theory and Practical Adequacy. In *Pragmatic Encroachment in Epistemology* eds. Brian Kim and Matthew McGrath. Routledge: New York.

[14] Condorcet, Le Marquis de. 1785. *Essai Sur L'Application de L'Analyse a la Probabilite Des Decisions Rendues a la Pluralite Des Voix*. Paris: L'Imprimerie Royale.

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

[16] Dorr, Cian, Jeremy Goodman, and John Hawthorne. 2014. Knowing Against the Odds. *Philosophical Studies*, 170, pp. 277-287.

[17] Fantl, Jeremy and Matthew McGrath. 2002. Evidence, Pragmatics, and Justification. *Philosophical Review*, 111, pp. 67-94.

[18] Fassio, Davide and Jie Gao. 2021. Do We Really Need a Knowledge-Based Decision Theory? *Synthese*, 199, pp. 7031-7059.

[19] Greaves, Hilary and William MacAskill. manuscript. The Case for Strong Longtermism.

[20] Goodman, Nelson. 1955. *Fact, Fiction, and Forecast*. Cambridge, Massachusetts: Harvard University Press.

[21] Goodman, Jeremy and Bernhard Salow. manuscript. Epistemology Normalized.

[22] Greco, Daniel. 2013. Probability and Prodigality. In *Oxford Studies in Epistemology, Volume 4* eds. Tamar S. Gendler and John Hawthorne. Oxford: Oxford University Press.

[23] Harsanyi, John C. 1985. Acceptance of Empirical Statements: A Bayesian Theory Without Cognitive Utilities. *Theory and Decision*, 18, pp. 1-30.

[24] Hawthorne, John and Amia Srinivasan. 2013. Disagreement without Transparency: Some Bleak Thoughts. In *The Epistemology of Disagreement: New Essays* eds. Jennifer Lackey and David Christensen. Oxford: Oxford University Press.

[25] Hawthorne, John and Jason Stanley. 2008. Knowledge and Action. *The Journal of Philosophy*, 105, pp. 571-590.

[26] Isaacs, Yoaav. 2016. Probabilities Cannot Be Rationally Neglected. *Mind*, 125, pp. 759-762.

[27] Kahneman, Daniel and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47, pp. 263-292.

[28] Kaplan, Mark. 1996. *Decision Theory as Philosophy*. Cambridge: Cambridge University.

[29] Lance, Mark N. 1995. Subjective Probability and Acceptance. *Philosophical Studies*, 77, pp. 147-179.

[30] Levi, Isaac. 1980. *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability and Chance*. Cambridge: MIT Press.

[31] Lewis, David. 1980. A Subjectivist's Guide to Objective Chance. In *Ifs* eds. William Harper, Robert Stalnaker, and Glenn Pearce. Springer, Dordrecht.

[32] Mueller, Andy and Jacob Ross. 2017. Knowledge Dethroned. *Analytic Philosophy*, 58, pp. 283-296.

[33] Monton, Bradley. 2019. How to Avoid Maximizing Expected Utility. *Philosophers' Imprint*, 19, pp. 2-25.

[34] Moss, Sarah. 2013. Epistemology Formalized. *Philosophical Review*, 122, pp. 1-43.

[35] Moss, Sarah. 2021. Knowledge and Legal Proof. In *Oxford Studies in Epistemology, Volume 7* eds. Tamar S. Gendler and John Hawthorne. Oxford: Oxford University Press.

[36] Nover, Harris and Alan Hájek. 2004. Vexing Expectations. *Mind*, 113, pp. 237-249.

[37] Parfit, Derek. 1981. Correspondence. *Philosophy and Public Affairs*, 10, pp. 180-181.

[38] Pryor, James. 2000. The Skeptic and the Dogmatist. *Noûs*, 34, pp. 517-549.

[39] Ross, Jacob and Mark Schroeder. 2014. Belief, Credence, and Pragmatic Encroachment. *Philosophy and Phenomenological Research*, 88, pp. 259-288.

[40] Rothschild, Daniel and Levi Spectre. 2018. A Puzzle about Knowing Conditionals. *Noûs*, 52, pp. 473-478.

[41] Schulz, Moritz. 2017. Decisions and Higher-Order Knowledge. *Noûs*, 51, pp. 463-483.

[42] Smith, Nicholas J. J. 2014. Is Evaluative Compositionality A Requirement of Rationality? *Mind*, 123, pp. 457-502.

[43] Smithies, Declan. 2012. The Normative Role of Knowledge. *Noûs*, 46, pp. 265-288.

[44] Spiess, Von O. 1975. Zur Vorgeschichte des Petersburger Problems. In *Die Werke von Jakob Bernoulli, Bd. 3: Wahrscheinlichkeitsrechnung* ed. Bartel Leendert van der Waerden. Basel: Birkhäuser.

[45] Staffel, Julia. 2019. How do Beliefs Simplify Reasoning? *Noûs*, 53, pp. 937-962.

[46] Stanley, Jason. 2005. *Knowledge and Practical Interests*. Oxford: Oxford University Press.

[47] Turri, John. 2017. Knowledge Attributions and Behavioral Predictions. *Cognitive Science*, 41, pp. 2253-2261.

[48] Weatherson, Brian. 2012. Knowledge, Bets and Interests. In *Knowledge Ascriptions* eds. Jessica Brown and Mikkel Gerken. Oxford: Oxford University Press.

[49] Wedgwood, Ralph. 2012. Outright Belief. *Dialectica*, 66, pp. 309-329.

[50] Weisberg, Jonathan. 2013. Knowledge in Action. *Philosophers' Imprint*, 13, pp. 1-23.

[51] White, Roger. 2006. Problems for Dogmatism. *Philosophical Studies*, 131, pp. 525-557.

[52] Wilkinson, Hayden. 2022. In Defence of Fanaticism. *Ethics*, 132, pp. 445-477.

[53] Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

[54] Williamson, Timothy. 2005. Contextualism, Subject-Sensitive Invariantism and Knowledge of Knowledge. *Philosophical Quarterly*, 55, pp. 213-235.

[55] Williamson, Timothy. 2007. How Probable is an Infinite Sequence of Heads? *Analysis*, 67, pp. 173-180.

[56] Wright, Crispin. 2004. Warrant for Nothing (and Foundations for Free)? *Aristotelian Society Supplementary*, 78, pp. 167-212.