You've Come a Long Way, Bayesians

Jonathan Weisberg
University of Toronto

Abstract. To celebrate the 40th anniversary of the *Journal of Philosophical Logic*, this article provides a retrospective on select topics from the last 40 years of Bayesian epistemology. Topics discussed include (1) scoring rules and accuracy arguments, (2) imprecise credences, (3) regularity and zero-probable events, (4) connections between Bayesianism and "informal" epistemology, and (5) full and partial belief.

Forty years ago, Bayesian philosophers were just catching a new wave of technical innovation, ushering in an era of scoring rules, imprecise credences, and infinitesimal probabilities. Meanwhile, down the hall, Gettier's (1963) paper was shaping a literature with little obvious interest in the formal programs of Reichenbach, Hempel, and Carnap, or their successors like Jeffrey, Levi, Skyrms, van Fraassen, and Lewis. And how Bayesians might accommodate the discourses of full belief and knowledge was but a glimmer in the eye of Isaac Levi.

Forty years later, scoring rules, imprecise credences, and infinitesimal probabilities are all the rage. And the formal and "informal" traditions are increasingly coming together as Bayesian arguments spill over into debates about the foundations of empirical knowledge, skepticism, and more. Relatedly, Bayesian interest in full belief and knowledge has never been greater.

Much more besides has happened in the last forty years of Bayesian philosophy, but there's far too much to cover it all here. So I've selected a grab-bag of topics. Some are hot—they are where the breaking news seems to be. Others are where my heart is, topics I'd like to think are heating up. All of them connect in interesting ways with where Bayesian philosophy, and epistemology more generally, stood forty years ago. Several have also been heavily influenced by the work of this journal's founding editor-in-chief. The result is partly a retrospective, partly a snapshot of the Bayesian moment, and partly a wishful thought about where things might go next.

This article is dedicated to the memory of Herbert Weisberg, for whom the last forty years were dedicated to family. Thanks to Kristen Aspevig, Kenny Easwaran, Franz Huber, and Alexander Pruss for much helpful discussion.

1. Scoring Rules & Accuracy Arguments

For epistemologists of a pragmatic bent, Dutch book arguments and representation theorems have traditionally secured Bayesianism's foundations. For epistemic purists, scoring rules now offer an alternative. Savage's (1971) and de Finetti's (1974) work on scoring rules laid the groundwork for a powerful, non-pragmatic approach to vindicating Bayesian epistemology. That approach hit the big time in (Joyce, 1998), a paper that has sparked an explosion of philosophical work on the strictly epistemic virtues of Bayesian epistemology.[1]

Savage and de Finetti proved that violating the probability axioms makes one's credences unnecessarily inaccurate: every credence function violating the probability axioms is *accuracy-dominated* by a probabilistic one, i.e. there is a credence function obeying the probability axioms that is more accurate—"closer to the truth"—*no matter what the truth turns out to be*.[2] Moreover, that probabilistic credence function is not accuracy-dominated by any other credence function.

Savage and de Finetti measured inaccuracy using *quadratic loss rules*, generalizations of Brier's (1950) famous scoring rule, according to which a probability's inaccuracy is the square of the difference between the probability and the proposition's truth-value (where truth is 1, falsity 0). The overall inaccuracy of a probability distribution in a "possible world" is the average of the individual scores over all propositions under consideration. Some propositions might be more epistemically important than others, but we can weight some terms in this sum more than others accordingly. The possible weightings generate the class of quadratic loss functions.

But why measure the inaccuracy of our credences using quadratic loss rules? Finding no convincing argument in the literature, Joyce (1998) set out to generalize Savage and de Finetti's result. He defended six constraints a reasonable measure of inaccuracy must obey, and then showed a similar result: every non-probabalistic credence function is always accuracy-dominated by some probabilistic one.

Maher (2002) objected that Joyce's arguments for two of his constraints were inadequate, and that they even ruled out some plausible scoring rules. For example, Joyce's constraints exclude the simple, linear scoring rule that just sums up the differences between credences and truth-values, rather than averaging the squares of differences. Joyce (2009) offers new arguments for the constraints in his (1998), and against the simple linear measure. But he simultaneously acknowledges that worries

[1] This is one area where the founding editor-in-chief of this journal did influential work: (van Fraassen, 1983).

[2] They gave this result a very different, pragmatic interpretation though.

raised by Maher and others (Gibbard, 2008; Hájek, 2009) make a new approach based on different assumptions desirable.

Joyce (2009) proposes two constraints on a reasonable scoring rule. The first is *Truth-Directedness*: given two credence functions b and b^* and a possible world w, if for every proposition A, $b(A)$ is at least as close to A's truth-value in w as $b^*(A)$ is, and there is at least one proposition where b's credence is closer than b^*'s, then b scores better than b^*. The second is *Coherent Admissibility*: our scoring rule should not let any probabilistic credence function turn out to be accuracy-dominated. If our scoring rule is finite, continuous, and satisfies these two constraints, and we restrict our attention to credences over a partition,[3] Joyce shows the desired result: every non-probabilistic credence function scores worse than some probabilistic one no matter how the world turns out to be. Moreover, the dominating probabilistic function is not dominated by any other credence function, an important result that wasn't shown in Joyce's (1998) (though Joyce affirms in personal communication that it nevertheless holds).

The second criterion here, Coherent Admissibility, appears to beg the question in favour of probabilistic credences. But Joyce argues it merely says that our scoring rule should not rule out any probabilistic credence functions a priori. And that's a far cry from ruling out non-probabilistic credence functions a priori.

Hajek (2009) and Leitgeb & Pettigrew (2010b) nevertheless worry that Coherent Admissibility assumes too much. As Hajek notes, an opponent who endorses non-probabilistic credences might likewise insist that no reasonable scoring rule should rule out her preferred functions a priori, leading to a stalemate. Joyce responds that probabilistic credences are special because an agent whose evidence said those probabilities were the objective chances would be rational to have those probabilities as her credences, in accord with the Principal Principle (Lewis, 1980). But, Hajek objects (and Leitgeb & Pettigrew concur), this won't work for propositions that don't have objective chances, or whose chance is always 0 or 1. (Examples might include laws of nature or chance-statements themselves.) A scoring rule that ruled out non-extreme probabilistic credences in such propositions a priori would not go against the Principal Principle, and so could not be dismissed on the grounds Joyce provides.[4,5]

[3] Joyce asserts that the restriction to a partition can be dropped, but does not prove this in the paper.

[4] I would add that an opponent of Probabilism might well reject the Principal Principle as standardly formulated. She might prefer to explicate the dictum that one's credences should reflect known chances other than by just having one's credences line up with those chances numerically.

[5] See (Easwaran & Fitelson, 2012; Fitelson, 2012; Carr, 2013, manuscript) for other criticisms of Joyce's overall approach.

Leitgeb & Pettigrew (2010a; 2010b) offer a different approach. They suggest that, more than merely avoiding accuracy-dominated credences, we should hold credences that minimize our *expected* inaccuracy. They argue on the basis of this norm that only quadratic loss functions are reasonable scoring rules, and then show that all and only probabilistic credence functions minimize expected inaccuracy when inaccuracy is measured by quadratic loss.

A worry for this approach is its reliance on the concept of expected value, which is intimately tied to probability theory. The additivity of probabilities means that expected values can be calculated using different partitions, hence different levels of fine-ness of grain, without affecting the result. But for non-additive functions, like the belief functions of Dempster-Shafer theory (Shafer, 1976), this invariance does not hold and the notion of "expected value" has no univocal extension. Leitgeb & Pettigrew (2010a, 214–5) address the problem by fixing on a single, privileged partition, namely the finest one: the set of singleton possibilities. But from the point of view of an alternative like Dempster-Shafer theory, this is entirely illegitimate. First, Dempster-Shafer theory allows all these atoms to have 0 credence, even when there are only finitely many.[6] This immediately takes Dempster-Shafer theory out of the scope of Leitgeb & Pettigrew's result, since they exclude such "zero-atomic" distributions by fiat. Second, even for credence functions that are not zero-atomic in Dempster-Shafer theory, weighting expectations only according to the credences in the atoms ignores information higher up in the "lattice of propositions". Dempster-Shafer theory allows one's credence in $\{w_1, w_2\}$ to be greater than the sum of one's credences in $\{w_1\}$ and $\{w_2\}$ individually. So using only the latter two credences to weight our expectation underweights one's credence in $\{w_1, w_2\}$, failing to give full credit for accuracy where credit is due.

Still, if we take the probability axioms as a given, we could then help ourselves to expected-inaccuracies for the purposes of justifying further Bayesian tenets. Greaves & Wallace (2006) argue that conditionalization is justified because it maximizes expected epistemic utility, and Leitgeb & Pettigrew (2010b) similarly offer an expected accuracy argument for conditionalization. Interestingly though, they find that accuracy norms come apart in cases where Jeffrey conditionalization is meant to apply. They conclude that insofar as an accuracy-based argument is available, it favours a different update-rule. In other applications, Pettigrew (2012) uses expected accuracy to justify various versions of the Principal Principle, though he returns to a Joycean, dominance-based approach in (Pettigrew, 2013). Pettigrew (forthcoming)

[6] This happens when the available evidence doesn't point to any individual possibility specifically, even though it perhaps point to pairs of them, triples, etc.

argues that the Principle of Indifference minimizes risk of inaccuracy, and Moss (2011) applies expected epistemic values to disagreements.

2. Indeterminate Credences

Bayesianism Classic represents an agent's credences with a probability function, assigning precise, real numbers to propositions. But some Bayesians now think rational opinion needn't be so determinate (Levi, 1974). Sometimes the evidence doesn't warrant any single, precise opinion, making a set of probability functions a more appropriate representation.[7] Following van Fraassen (1990), the set of probability functions representing an indeterminate state of opinion is called the agent's *representor*, and her credence in a particular proposition is capture by a set of values, typically an interval.

There's been some intense debate about this approach recently. Objectors worry that it faces serious problems with updating and decision-making. White (2010) presses the challenge for updating, Elga (2010) the challenge for decision-making. We'll focus on decision-making here.

How should an agent with indeterminate credences make choices? One natural proposal is to allow any choice that maximizes expected utility with respect to some member of her representor.[8] Suppose your credence in H is the interval $[1/10, 8/10]$ and you are offered the following bet to accept or decline as you like:

Bet A: Lose $10 if H, win $15 otherwise.

On the current proposal you may either accept or reject the bet, since some members of your representor give it an expected gain while others rate its expectation zero or negative.

Trouble is, this permissive policy can start to look foolish over time. Suppose you decline Bet A as permitted, and you're immediately offered another bet:

Bet B: Win $15 if H, lose $10 otherwise.

As before, you're permitted to accept or decline. But if you do decline, you've effectively missed out on a free $5. Had you taken both bets, you would have lost $10 on one but won $15 on the other, for a net gain of $5 no matter what. Even if you

[7] Authors who base their Bayesianism on representation theorems may have other motivations for embracing imprecise credences (Jeffrey, 1987; Joyce, 1999).

[8] Other proposals are on offer. Elga groups these under various rubrics and argues that each group only escapes his argument at unacceptable cost. Chandler (forthcoming) replies that one proposal, the Γ-maximin rule, satisfies an appealing standard that is slightly weaker than the one Elga's argument assumes. Bradley & Steele (2014) argue that Γ-maximin nevertheless faces other problems raised in (Seidenfeld, 2004) and (Steele, 2010).

knew ahead of time that you would be offered these two bets in sequence, the present proposal would still permit you to make choices that guarantee you $0 when you could guarantee yourself $5.

Could we respond that your first choice commits you to a more precise credence, and thereby constrains your second choice? If you decline A because some members of your representor assign it zero or negative expectation, maybe then you're committed to a new representor with only those members, in which case you'll find that accepting B is compulsory. But Elga counters that you would then be changing your opinion without relevant evidence, since your choice to reject Bet A tells you nothing about whether H is true.

Could we respond instead that a rational agent faced with this sequence of offerings would form a plan about how to proceed? Any rational plan would have to involve taking at least one of A or B, so following a plan ensures that you won't leave $5 sitting on the table. Here Elga objects that irrelevant factors would have to be introduced into our decision theory. Suppose your plan is to reject A and accept B. Having rejected A, must you now accept B? If you had only ever been offered B and thus had no plan, you'd be permitted to reject B. And all relevant factors look to be the same in this alternate scenario. Your relevant evidence, beliefs, desires, and goals are all the same. That you have been planning to take B doesn't make B any more effective a means to promoting your interests. So, since you would be reasonable to decline B if you had no plan, you would still be reasonable to decline B despite having planned to accept it.

Joyce (2010) responds that, despite initial appearances, the imprecise approach does not permit rejecting both A and B. While you are not obliged to accept or to reject A, and likewise for B, you are obliged to accept B *if you reject A*. The reason is that any probability function in your representor that licenses rejecting A will oblige you to accept B. To paraphrase Joyce: it is a determinate fact about you that you prefer to reject A only if you prefer to accept B. This is so despite there being no determinate fact about whether you prefer to accept/reject A, and likewise for B.

But this determinate, conditional fact about what you prefer is not enough to stop you from realizing it in two different ways at two different times. At t_1 you might prefer to reject A and accept B, but then at t_2 there is nothing to stop you reversing these preferences, since either way is consistent with your credences and utilities. So now you may prefer to reject B and to accept A, with the end result that you reject both. To make matters worse, choosing to reject A at t_1 needn't even reflect any preference on your part. By hypothesis, you have no preference between accepting or rejecting A. So you may choose to reject A simply because you must choose something. But this says nothing about what you prefer, and thus activates no modus ponens involving the determinate, conditional fact that you prefer to accept B if you prefer to reject A.

Joyce also suggests that agents faced with incommensurable choices can resolve their indecision using "pseudo-beliefs," obtained by sharpening their representor towards a midpoint. Faced with the choice about A, for example, you could sharpen your credence in H towards the midpoint 45/100, arriving at the "pseudo-belief" $(4/10, 5/10)$. You would then reject A and, using the same recipe later, accept B. Joyce is careful to insist you not change your actual credences here, since then Elga would object that you are changing your opinion based on irrelevant information. Rather, you merely use the constructed pseudo-beliefs as if they were your real credences for the purely pragmatic purpose of resolving indecision.

But if pseudo-beliefs are constructed purely for the pragmatic end of resolving indecision, why should they persist from one decision to the next? Symmetrical narrowing is just one way of resolving indecision, and Joyce explicitly allows that one can resolve indecision in different ways in different contexts. So why can't one sharpen towards the midpoint when deciding to reject A, but later resolve one's indecision about B in a different way, one that recommends rejecting B? If the rationale behind sharpening is just the pragmatic one that it serves the aim of reaching a decision, it is just as arbitrary to sharpen towards the midpoint each time as to use different methods for resolving indecision each time.

Moss (forthcoming) offers a different approach, where agents with imprecise credences are viewed as torn between the different states of mind contained in their representors, much as an agent facing a moral dilemma might be represented by a set of conflicting utility functions. When faced with the decision whether to accept or reject an option like Bet A, an agent faces a "credal dilemma", which she resolves by "identifying with" one of the precise states of mind in her representor. Because she will not normally change her mind to identify with a different element in her representor before facing the decision about B, she will typically accept B if she rejects A.

Moss does want to allow that an agent can identify with different elements of her representor at different times (it would be hard to see why they count as members of her representor otherwise). But for her to change her identification, she must undergo an honest change of mind, and in that case we would not blame her for leaving $5 on the table. For example, Moss describes an agent whose evidence does not settle which course of action is more likely to lead to the desired outcome, leaving her torn between two states of mind. The agent makes a decision, but after sleeping on it she wakes with a pit in her stomach and changes her mind. Despite the fact that such an agent might leave $5 on the table, she is rational according to Moss, since her actions result from a genuine and reasonable change of mind.

A worry for Moss' approach is whether it is general enough. It's plausible to think of an agent as torn between competing states of mind in cases where the evidence lends itself to competing lines of thought. And Elga works with such a case in his paper. But the literature on imprecise credences is inspired as much by cases where the evidence is *incomplete* as by those where it is *conflicting*. The stock examples are urns full of marbles mixed in unknown proportions, like in the Ellsberg paradox (Ellsberg, 1961). There an urn contains 90 marbles, 30 of which are known to be red, the rest either black or yellow but in unknown proportions. Here it's harder to see the agent as "identifying with" any one element in her representor. In cases of conflicting evidence, we might find ourselves more in the grip of one line of reasoning than another, or inclined to give one line of thought more weight. But in the Ellsberg case there is no line of argument drawing us towards any one element of the representor—all are on a par. So Moss' assumption that agents resolve credal dilemmas by "identifying with" a sharpening is harder to accept here. Joyce's thought that the agent picks a pseudo-belief for the purely pragmatic purpose of resolving indecision seems more descriptively accurate, but this brings us back to Elga's challenge.

I prefer a divide and conquer strategy. Elga describes his argument as a Dutch book variant. Let's first divide by recalling a stock objection-and-reply for such arguments. The objection: a Dutch book argument shows at best that the agent is *prudentially* irrational, not that she is *epistemically* irrational. The reply: the prudential folly just serves to dramatize an epistemic folly.[9] In the case of the Dutch book argument for additivity, for example, we dramatize an inconsistency in the agent's commitments, since she is committed (supposedly) to regarding as fair a set of bets that is clearly not. But in Elga's example, choosing to reject A and to reject B cannot expose any such inconsistency in your commitments, for your choices do not reflect *any* commitment. Your credences and utilities underdetermine whether each bet is fair, since they are consistent with either assessment. For each bet, that it is fair is consistent with your beliefs and utilities, but so is the assessment that it is unfair. What Elga's example exposes is not that your beliefs are inconsistent, but that they are *incomplete*. And there is nothing irrational about an incomplete state of opinion. Suspending judgment is frequently the rational thing to do.[10]

[9] See (Vineberg, 2014) for more on this dialectic.

[10] Consider what Elga's argument would prove about an agent only capable of on/off beliefs, whose evidence justified neither believing nor rejecting H.

Perhaps Elga only means to show that imprecise credences are *prudentially* irrational? If so, we might just accept that epistemic and prudential norms conflict here. Perhaps epistemic normativity sometimes requires imprecise credences despite prudential normativity requiring precise ones. Respecting the evidence might require one thing, pursuing one's ends another. But we needn't even concede that much.

First because Elga's central assumption is harder to motivate from a strictly prudential point of view. Why is it prudentially irrational to reject A and then reject B? It would be irrational to reject both A and B were they offered together. That would be a strictly dominated choice. But an agent with imprecise credences would not be allowed to reject both simultaneously. And Elga's setup is quite different. As Bradley & Steele (2014) note, when you reject A you are not making a decision that leads to a sure loss (or the loss of a sure gain), since you may yet accept B. And if you then reject B, you are again not inviting the loss of a sure gain since accepting B could cost you \$10. Elga might argue that if each of two actions is prudentially rational individually, they must also be prudentially rational as a package. But the package deal in Elga's case is neither rational nor irrational—it is not within the purview of our decision theory, for it is not an available option (cf. Chandler, forthcoming, §6). When the two bets are offered in sequence, there is no single action that corresponds to rejecting both (unless the agent can bind herself to a plan, in which case the puzzle dissolves).

Second, even if we accepted Elga's central assumption, we could just insist that you ought to accept B once you reject A. After all, this is consistent with your credences and utilities, since they fail to adjudicate between accepting and rejecting B. Elga may object that we then allow irrelevant factors to influence your decision, since your acceptance of A makes B no better as a prospect. But if your present credences and desires are all that matter to the rationality of your accepting B, Elga cannot insist that your choices on A and B are rationally bound up. He cannot have it both ways, insisting that the two choices are a package deal from the point of view of prudential rationality, yet each must be made independently of the other.

3. Improbability & Impossibility

Should a live epistemic possibility ever get zero credence? There are several reasons to think not, here are three. First, zero-probable propositions cannot be conditionalized on in classical probability theory, since the requisite conditional probabilities will be undefined: $p(A|B) =_{\text{df}} p(A \wedge B)/p(B)$, which is undefined when $p(B) = 0$. Second, a zero-probable proposition can never become more probable by conditionalizing on new information, since $p(A|B) = 0$ whenever $p(A) = 0$. Third, zero-probable possibilities cannot influence expected utilities, resulting in no incentive to take the free offer of a million dollars if some zero-probable possibility obtains.[11]

On the other hand, in standard probability theory it's impossible to avoid assigning probability zero to some possibilities if there are uncountably many of them. Even in the countable case, the only way for a lottery with a countable number of tickets to be fair is for each ticket to have zero probability. Anything larger, and the probability of some finite subset of the tickets will add up to more than one. Assigning zero to each ticket means abandoning countable additivity, as de Finetti (1937 [1980]) advocates, so we might instead conclude that no countable lottery could be fair. But in the uncountable case, even abandoning fairness won't help. A real-valued probability assignment over an uncountable space of possibilities must assign uncountably many zeros (Williamson, 2007).

The natural move then is to abandon the classical framework for a non-standard one, expanding the range of probability-values to include infinitesimals: probabilities greater than zero but less than any positive real number (Skyrms, 1980; Lewis, 1980, 1986). Then, not only can we avoid assigning probability zero when faced with infinite possibilities, but we can even do so in uniform and fair ways (Bernstein & Wattenberg, 1969). We can even preserve the idea that the probability of the whole should be the sum (in some sense) of the probabilities of the parts, at least for countable possibility spaces (Wenmackers & Horsten, 2013).

The relief may only be temporary though. Williamson (2007) argues that, even if we allow infinitesimal probabilities, some possibilities will still have probability zero. For suppose we are about to toss a fair coin infinitely many times, with each toss independent of the others. The probability that all the tosses will come up heads is twice the probability that at least all tosses after the first will be heads. Intuitively this is because the former proposition takes all the same risks as the latter, plus the additional risk of a 50% chance of being wrong about the first toss coming up heads. More rigorously, let H_n be the proposition that the n-th toss is heads, and $H_n\ldots$ the proposition that the n-th toss and on are all heads. Then:

$$\begin{aligned} p(H_1\ldots) &= p(H_1 \wedge H_2\ldots) && (1)\\ &= p(H_1)p(H_2\ldots|H_1) && (2)\\ &= (1/2)p(H_2\ldots|H_1) && (3)\\ &= (1/2)p(H_2\ldots). && (4)\end{aligned}$$

[11]See (Hájek, manuscript, §9) for a fuller catalogue of motivations.

The first two lines follow by probability calculus, the latter two from our fairness and independence assumptions. And yet the probabilities of $H_1 \ldots$ and $H_2 \ldots$ are the same, since these are isomorphic, qualitatively identical events:

$$p(H_1 \ldots) \quad = \quad p(H_2 \ldots). \tag{5}$$

So $p(H_1 \ldots)$ must be zero!

There aren't a lot of outs here. We could say that no such sequence of tosses is possible, or that no rational agent can believe it is as described, but no cause for either claim is forthcoming. So (1)–(4) look pretty unassailable. What about (5)? Weintraub (2008) objects that the events described by $H_1 \ldots$ and $H_2 \ldots$ aren't really isomorphic, since one has a property the other lacks. Only the later sequence of heads is a proper subset of the first. But this extrinsic, relational difference doesn't seem relevant to the kind of physical isomorphism Williamson cites in support of (5).

Perhaps better to allow that these events are physically isomorphic, but say that physical isomorphism does not always mean sameness of probability. Ordinarily the (physical) probability of an event supervenes on intrinsic physical features like those shared by these two isomorphic events. But just as we ordinarily think that whether there is room in the hotel is determined by how many unoccupied rooms there are, yet find ourselves mistaken in the infinite case, so too here. After all, from the point of view of someone who antecedently thinks $H_1 \ldots$ and $H_2 \ldots$ have non-zero probability, (5) could only be true if there were no risk in logically strengthening $H_2 \ldots$ by conjoining H_1, i.e. only if H_1 had probability 1, contra our assumptions.

Even so, ugliness emerges if we embrace infinitesimal probabilities, since the resulting conditional probabilities can be very strange. Pruss (2012) adapts a result from (Dubins, 1975) to highlight just how troubling this strangeness can be: I can convince you of any proposition I like just by hitching the right kind of random process to its truth value. To illustrate, suppose I want to convince you the earth is a cube, which you rate vastly improbable, say $1/10^{10}$. Still, I can make you virtually certain it's a cube as follows. First I determine whether the earth is a cube, though I don't tell you my finding (you may be tempted to guess, but wait). Instead I inform you that if the earth is not a cube, I will choose a positive integer n using a uniform, random process, and report that number to you. If it is a cube, I'll instead choose n by a process that has probability $1/2^n$ of choosing n, and report that result. You won't see which process generates my report, but whatever I report, you will become certain the earth is a cube.

Why? Because any given report is infinitely more likely if the earth is a cube than if it's not. The probability that I'll report (say) '$n = 101$' is $1/2^{101}$ if the earth is a cube, while it's infinitesimally small if the earth is not a cube. So I was infinitely more likely to make that report if the earth is a cube. Formally:

$$\begin{aligned} p(Cube \mid n = 101) &= \frac{p(n = 101 \mid Cube)p(Cube)}{p(n = 101 \mid Cube)p(Cube) + p(n = 101 \mid \neg Cube)p(\neg Cube)} \\ &= \frac{(1/2^{101})(1/10^{10})}{(1/2^{101})(1/10^{10}) + \varepsilon(1 - 1/10^{10})} \\ &= \frac{1}{1 + \varepsilon(10^{10} - 1)(2^{101})} \end{aligned}$$

where ε is the infinitesimal probability that I will choose a given number if I use a uniform, random process. Of course, the same goes for any other number I might report. The cube hypothesis always fits the evidence better than its negation—infinitely better!

You can see this absurd result coming, which highlights the bizarre feature of infinitesimal distributions responsible: they are not conglomerable.[12] Thus they can lead to unacceptable violations of the Reflection Principle (van Fraassen, 1984). While there are many cases where it is reasonable to violate Reflection, it is not reasonable when you know your opinion will only change by conditionalizing on additional evidence (Weisberg, 2007; Briggs, 2009).

Pruss (2013) has still more bad news. Admitting infinitesimals into the range of our probability function only enables us to avoid zeros on countable or continuum-sized domains. Larger domains will always need even larger ranges to avoid zeros. Very crudely the reason is this: if every possible world has positive probability, then adding a world to a set always increases its probability. Thus we'll need distinct probability values for each of $\varnothing$, $\{w_1\}$, $\{w_1, w_2\}$, $\{w_1, w_2, w_3\}$, … until we run out of worlds. So the more worlds there are, the more probability-values we'll need available. Pruss shows that, given quite weak assumptions, this reasoning generalizes nastily beyond the finite case. If the domain of possibility is larger than the probability function's range, some possibility must have probability zero.[13]

Bayesians might take some comfort in the fact that some agents' epistemic spaces will be limited to the size of some cardinal. For these agents, some extension of the standard axioms of probability theory might be formulated. But these agents will not be able to formulate the Bayesian theory of rationality that applies to them, and

[12] That is, a proposition's unconditional probability can lie outside the span of its conditional probabilities across a partition.

[13] See (Hájek, manuscript) for a different argument in this "arms-race" style.

no formulation exists for agents who (like readers of this journal) are aware of the entire ZFC hierarchy. The best they might hope to do is formulate theories for more naive agents to use as approximations.

Well, if we must give zero credence to some live possibilities, can we at least avoid the troubling consequences that opened this section? We can replace Kolmogorov's probability axioms with one's that, like Popper's (1959) or Rényi's (1970), begin with conditional probability rather than unconditional probability as the primitive notion. Then conditional probabilities can be defined even when the condition has probability zero.[14] And conditionalizing on zero-probable conditions can raise the probability of zero-probable events. The decision-theoretic problem remains unsolved, however.

4. "Informal" Bayesianism

Formal approaches to questions of "informal" epistemology are becoming increasingly common. Bayesian analyses are now frequently brought to bear on debates about internalism vs. externalism, skepticism, and coherentist vs. foundationalist theories of knowledge. We'll look at a few examples here.

Must we know that our perceptual faculties are reliable before we can trust them? If so, the regress problem threatens us with skepticism. To know there is a glass of water to my left, as there appears to be, I must first know that my vision is reliable. But knowledge of my vision's reliability then needs justification from some other source, raising the same problem again, ad infinitum. Some epistemologists respond by embracing the notion of *immediate justification*: a perception as of P justifies (defeasibly) the belief that P all by itself, without any aid from, or need for, antecedent knowledge that perception is reliable. (Pollock, 1971, 1974, 1995; Pryor, 2000, 2005)

White (2006) argues that Bayesian considerations bode ill for the possibility of immediate justification. Let E be the perceptual evidence that there appears to be a glass of water to my left, and H the hypothesis that there really is a glass there. Then $E \wedge \neg H$ is the possibility that this perceptual evidence is misleading, that the appearance is not the reality. As a matter of probability calculus, $p(H|E) \leq p(\neg(E \wedge \neg H))$. So the perceptual evidence cannot make me confident of H unless I was already confident that perception would not mislead me here. Apparently perceptual justification is not immediate, but instead depends on antecedent knowledge (or justified belief) that perception is reliable.[15]

On the other hand, Bayesianism seems to get the interaction between perception and background information wrong anyway, whether we embrace immediate

[14] See (Fitelson & Hájek, manuscript) for some additional benefits for our definition of independence.
[15] Pryor (2013) highlights some assumptions tacit in this argument.

justification or not. Even if I do know that my perceptual faculties are reliable, updating by (Jeffrey) conditionalization leads to the absurd result that I won't abandon perception-based beliefs even after learning that perception was unreliable in the circumstances. (Weisberg, 2009, forthcoming; Pryor, 2013) Wagner (2013) argues that this problem can be avoided if we are more judicious in choosing the proposition we update on, e.g. by updating on the material conditional, *Perception is reliable in the circumstances* ⊃ *There is a glass to my left*. But as Gallow (2014) observes, this has the unpalatable effect of increasing my confidence that perception is deceptive in the circumstances.

Gallow suggests instead that we replace conditionalization with a rule that takes account of background beliefs about whether the environment is deceptive or our faculties impaired. Gallow proposes such a rule, though its implications for immediate justification have yet to be explored in the literature. If Gallow's proposal proves correct, I expect White's argument against immediate justification will still be available. Gallow's rule, like conditionalization, makes your antecedent credence that your faculties are reliable a limiting factor on E's support for H, though arguing for this claim requires a more detailed discussion.

Bayesian analyses have also been brought to bear on Moorean responses to skepticism. If the appearance as of a hand justifies me in believing that I am perceiving a real hand, can it also justify me in believing that this apparent hand is not an illusion? After all, that it is not an illusion deductively follows from the fact that I am perceiving a real hand.

Here again, White (2006), and Silins (2008), wield a Bayesian counter-argument. Begin by observing that evidence for a hypothesis always increases the probability that the evidence is misleading. For suppose we get evidence E supporting some hypothesis H. Then the conjunction $E \wedge \neg H$ represents the possibility that the evidence is misleading. This conjunction entails E, so by the probability calculus, $p(E \wedge \neg H | E) > p(E \wedge \neg H)$ as long as $p(E) < 1$ (as is the case for new evidence). But it seems absurd to say that E justifies rejecting $E \wedge \neg H$ when it *increases* its probability.

A similar Bayesian analysis applies to the bootstrapping problem.[16] Suppose I know my local newspaper is reliable enough to be trustworthy, maybe even more reliable. One day I open the paper and read that P, thereby coming to know P. I also know the newspaper says P, so I deduce that the newspaper was correct in this instance. On this basis, I become very slightly more confident that the newspaper

[16] The bootstrapping problem first arose as a challenge for reliabilist theories of knowledge (Fumerton, 1995; Vogel, 2000), but as we see here it is generalizable into a paradox that makes trouble much more broadly (Neta, 2008; Weisberg, 2010, 2012).

is actually more reliable than I thought. My basis is admittedly slim ($N = 1$), but suppose my increase in confidence is correspondingly slim. Something has clearly still gone wrong here. And if I repeat this exercise a few more times, the wrongness gets amplified. There may be a limit to how many times I can repeat the exercise before I should begin to worry that my data set contains a bad data-point, since for all I know the newspaper does make occasional errors. But clearly something goes wrong well before I ever get to that point. What has gone wrong?

You might think I'm just guilty of circular reasoning (Vogel, 2008), but White (2006) observes that such reasoning is often fine. I can rely on my memory of a memory test to bolster my confidence in my memory's reliability. Maybe the problem is that my reasoning takes no epistemic "risks". It will vindicate the hypothesis that the newspaper is reliable no matter what I read (Titelbaum, 2010; Douven & Kelp, 2013). But here too we seem to find counterexamples, cases where "no-risk" reasoning looks okay (Vogel, 2000). Moreover, the thought that riskless reasoning can make no epistemic gains is normally underwritten by a theorem of probability that does not apply to the newspaper case (Weisberg, 2012).

A different probabilistic analysis does apply, however. As in the Moorean argument, I am trying to infer from a justified belief something which my basis for that belief does not probabilistically support. In this case, E is *The newspaper says P*, H is P, and I am trying to conclude from H (together with E) that the newspaper is more trustworthy than I thought. But that the newspaper says P is, by itself, probabilistically irrelevant to the newspaper's trustworthiness. (Weisberg, 2010, 2012) There are two notable differences between this case and the Moorean one. First, here my final inferential step is inductive while it was deductive in the Moorean case. Second, in the Moorean argument the final inferential step appeals only to H, while here it appeals to H together with E. But a simple Bayesian theorem supports the diagnosis here as there. We can easily prove that E can probabilify H, and $E \wedge H$ can probabilify H', without E probabilifiying H'.[17]

5. Full & Partial Belief

We just saw some interactions between Bayesianism and epistemologies concerned primarily with knowledge and full belief (as opposed to degrees of belief). What is the general relationship between degrees of belief and full beliefs? One issue here is the descriptive, metaphysical question whether full belief reduces to partial belief, or vice versa. Kenny Easwaran's entry in this volume (Easwaran, this volume) discusses that question, so we'll focus on other, normative questions. Should your partial beliefs

[17] See Douven (2011) for a general study of the intransitivity of probabilistic support.

constrain or affect your full beliefs? Vice versa? Should your actions be determined by your full beliefs, by your partial beliefs, or somehow by both?

One natural proposal is that you ought to fully believe P if you ought to have high credence in P.[18] But the lottery paradox presents an immediate problem: in a large, fair lottery with exactly one winner, you should believe that each ticket will lose, and hence that all will lose, even though it is stipulated that one ticket must win. Some respond by placing restrictions on when high credence warrants full belief, to exclude lottery cases (Pollock, 1995; Ryan, 1996; Douven, 2002). Others block the conclusion that all tickets will lose by rejecting the conjunction principle, that belief in P and belief in Q warrant belief in $P \wedge Q$ (Kyburg, 1970; Foley, 1993, 2009; Sturgeon, 2008).

The first approach must contend with Korb's (1992) observation that any belief can be "lotterized". By cleverly partitioning a believed proposition, we can make it a member of a set with the logical and probabilistic structure of a lottery. Let P be the proposition that I have hands, for example. Given a large, fair lottery with exactly one winner, P can be partitioned thus: $P \wedge$ *Ticket #1 will win*, $P \wedge$ *Ticket #2 will win*, etc. The propositions of the form $\neg(P \wedge$ *Ticket #i will win*$)$ form, together with P, the usual sort of problematic lottery: exactly one of them is false and each has the same high probability (or very nearly the same). If we say that high credence warrants full belief except when the logical and probabilistic structure of a lottery obtains, then my high credence that I have hands doesn't warrant full belief.

Korb's challenge is strengthened by Douven & Williamson (2006), who show that it generalizes to a broad class of restrictions one might propose, which includes the proposals of (Pollock, 1995), (Ryan, 1996), and (Douven, 2002). Their generalization has been criticized for assuming a finite set of possible worlds, each with the same probability (Chandler, 2010; Smith, 2010; Easwaran, this volume). But these assumptions may not weaken the argument as much as it seems. It is unlikely that your warrant for believing you have hands depends on being able to contemplate an infinitude of possibilities. And we may suppose that you have just learned about the existence of some lottery hitherto unconceived, leading you to refine your initial space of possible outcomes into one where all the possibilities are equiprobable.

What about the second approach, rejecting the conjunction principle, that belief in P and belief in Q warrant belief in $P \wedge Q$? The preface paradox (Makinson, 1965) suggests that skepticism and radical immodesty are the only alternatives anyway.

[18] Another promising approach we haven't space to explore: you ought to believe P if (roughly) doing so maximizes expected *epistemic* utility (Levi, 1967, 1980; Kaplan, 1996; Maher, 1993; Frankish, 2009).

Yet much of the lottery paradox would still remain unsolved. We avoid believing a contradiction, but we still end up believing that each lottery ticket will lose, which is unsettling on its own. The problem with lottery beliefs is not just that they are inconsistent (that happens in the preface paradox too). Lottery beliefs are additionally repugnant—maybe because their basis is purely statistical (Nelkin, 2000), maybe because they do not track the truth (Roush, 2005), or maybe because they do not constitute knowledge (Williamson, 2000). It is difficult to settle why lottery beliefs feel unwarranted, but it seems clear enough that they do.

What about normative questions going the other direction: do full beliefs contribute anything to the normative theory of partial beliefs? Levi (1980) suggests that an agent's full beliefs determine her "serious possibilities," which impacts her decision-making since only serious possibilities figure into her expected utilities. Foley (1993; 2009) endorses a similar view, that decision problems are shaped by our beliefs about the available acts, the possible states, and the resultant outcomes.[19] Foley also suggests that full beliefs shape degrees of belief. Your 1/13 credence that the next card drawn from a deck will be an ace is based on your full belief that the deck is standard and shuffled. In (Weisberg, 2013) I argue that empirical work in the psychology of judgment and decision supports a similar view. Much research suggests that we reason using methods like "Evidence Accumulation" (Lee & Cummins, 2004; Newell & Lee, 2011), which combine both graded and categorical judgments. Using these methods simplifies cognitive tasks that would be too onerous or wasteful to perform in pure Bayesian fashion.[20]

This view raises exciting but difficult questions. When are we permitted to simplify a cognitive task? And when these simplifications lead to choices or judgments that differ from what pure Bayesian reasoning would recommend, are these choices and judgments rational? I am inclined to think they are, at least if it was rational to use the simplified method in the first place. Whether it is rational to simplify in a given situation, and what simplifications are appropriate, depend on what practical and epistemic goods are at stake and what resources (time, energy, memory) the agent can spare.

A striking consequence of this view is that practical considerations can end up affecting what judgments are epistemically rational, resulting in rational differences of opinion even between agents with identical evidence. Some argue that such "epistemic permissiveness" is unacceptably arbitrary. If different credences can be

[19] Cf. also Weatherson (2012) and Ross & Schroeder (2014), though Weatherson argues that a decision table should only be shaped by beliefs constituting knowledge.

[20] See (Buchak, 2014) for a very different, novel view of the respective roles of full and partial belief.

warranted by the same evidence, you would be equally rational to hold one as you would be to hold the other, in which case you might as well flip a coin to decide what to believe (White, 2005). One attraction of the present proposal is that it answers this objection while also respecting the thought that reasonable people can draw different conclusions from the same evidence. The cause of the difference in opinion is not any arbitrary decision or self-serving bias. Rather, differences in circumstances and non-evidential resources like time and memory result in different methods of reasoning, even different executions of the same method.

Though speculative, these remarks suggest that a deeper appreciation of the interactions between Bayesian epistemology, the epistemology of full belief, and the psychology of judgment might produce fruitful insights into the big old questions of epistemology.

References

Bernstein, Allen R. & Frank Wattenberg. 1969. Non-standard measure theory. In W. A. J. Luxemberg (ed.), *Applications of model theory of algebra, analysis, and probability*, Holt, Rinehart and Winston: New York.

Bradley, Seamus & Katie Steele. 2014. Should subjective probabilities be sharp? *Episteme* 11(3): 277–289.

Brier, Glenn W. 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 78(1): 1–3.

Briggs, Rachael Amy. 2009. Distorted reflection. *The Philosophical Review* 118(1): 59–85.

Buchak, Lara. 2014. Belief, credence, and norms. *Philosophical Studies* 169(2): 285–311.

Carr, Jennifer. 2013. *Justifying bayesianism*: Massachusetts Institute of Technology dissertation.

Carr, Jennifer. manuscript. Epistemic utility theory and the aim of belief .

Chandler, Jake. 2010. The lottery paradox generalized? *British Journal for the Philosophy of Science* 61(3): 667–679.

Chandler, Jake. forthcoming. Subjective probabilities need not be sharp. *Erkenntnis* .

de Finetti, Bruno. 1937 [1980]. Foresight, its logical laws, its subjective sources. In Henry E. Kyburg & Howard E. Smokler (eds.), *Studies in subjective probability*, Krieger: Huntington, NY 2nd edn.

de Finetti, Bruno. 1974. *Theory of probability*. John Wiley and Sons.

Douven, Igor. 2002. A new solution to the paradoxes of rational acceptability. *British Journal for the Philosophy of Science* 53(3): 391–410.

Douven, Igor. 2011. Further results on the intransitivity of evidential support. *The Review of Symbolic Logic* 4(4): 487–97.

Douven, Igor & Christoph Kelp. 2013. Proper bootstrapping. *Synthese* 190(1): 171–185.

Douven, Igor & Timothy Williamson. 2006. Generalizing the lottery paradox. *British Journal for the Philosophy of Science* 57(4): 755–779.

Dubins, Lester E. 1975. Finitely additive conditional probabilities, conglomerability and disintegrations. *Annals of Probability* 3(1): 89–99.

Easwaran, Kenny. this volume. Formal epistemology. *Journal of Philosophical Logic* .

Easwaran, Kenny & Branden Fitelson. 2012. An 'evidentialist' worry about joyce's argument for probabilism. *Dialectica* 66(3): 425–433.

Elga, Adam. 2010. Subjective probabilities should be sharp. *Philosophers' Imprint* 10(5): 1–11.

Ellsberg, Daniel. 1961. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics* 75(4): 643–649.

Fitelson, Branden. 2012. Joyce's argument for probabilism. *Philosophy of Science* 79(1): 167–174.

Fitelson, Branden & Alan Hájek. manuscript. Declarations of independence .

Foley, Richard. 1993. *Working without a net: A study of egocentric epistemology*. Oxford University Press.

Foley, Richard. 2009. Beliefs, degrees of belief, and the lockean thesis. In Franz Huber & Christoph Schmidt-Petri (eds.), *Degrees of belief*, vol. 342 Synthese Library, 37–47. Springer.

Frankish, Keith. 2009. Partial belief and flat-out belief. In Franz Huber & Christoph Schmidt-Petri (eds.), *Degrees of belief*, vol. 342 Synthese Library, Springer.

Fumerton, Richard. 1995. *Metaepistemology and skepticism*. Rowman & Littlefield.

Gallow, J. Dmitri. 2014. How to learn from theory-dependent evidence; or commutativity and holism: A solution for conditionalizers. *British Journal for the Philosophy of Science* 65(3): 493–519.

Gettier, Edmund L. 1963. Is justified true belief knowledge. *Analysis* 23: 121–3.

Gibbard, Alan. 2008. Rational credence and the value of truth. In Tamar Szabo Gendler & John Hawthorne (eds.), *Oxford studies in epistemology*, vol. 2, Oxford University Press.

Greaves, Hilary & David Wallace. 2006. Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind* 115: 607–632.

Hájek, Alan. 2009. Arguments for—or against—probabilism? In Franz Huber & Christoph Schmidt-Petri (eds.), *Degrees of belief*, 229–252. Springer.

Hájek, Alan. manuscript. Staying regular .

Jeffrey, Richard C. 1987. Indefinite probability judgment: A reply to levi. *Philosophy of Science* 54(4): 586–591.

Joyce, James. 1998. A nonpragmatic vindication of probabilism. *Philosophy of Science* 65(4): 575–603.

Joyce, James. 1999. *The foundations of causal decision theory*. Cambridge University Press.

Joyce, James. 2009. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In Franz Huber & Christoph Schmidt-Petri (eds.), *Degrees of belief*, vol. 342 Synthese Library, 263–297. Synthese.

Joyce, James. 2010. A defense of imprecise credences in inference and decision making. *Philosophical Perspectives* 24(1): 281–323.

Kaplan, Mark. 1996. *Decision theory as philosophy*. Cambridge University Press.

Korb, Kevin B. 1992. The collapse of collective defeat: Lessons from the lottery paradox. In *Psa: Proceedings of the biennial meeting of the philosophy of science association*, vol. 1, 230–236.

Kyburg, Henry E. 1970. Conjunctivitis. In Marshall Swain (ed.), *Induction, acceptance, and rational belief*, D. Reidel.

Lee, Michael D. & Tarrant R. Cummins. 2004. Evidence accumulation in decision making: Unifying the "take the best" and the "rational" models. *Psychonomic Bulletin & Review* 11(2): 343–52.

Leitgeb, Hannes & Richard Pettigrew. 2010a. An objective justification of bayesianism i: Measuring inaccuracy. *Philosophy of Science* 77(2): 201–235.

Leitgeb, Hannes & Richard Pettigrew. 2010b. An objective justification of bayesianism ii: The consequences of minimizing inaccuracy. *Philosophy of Science* 77(2): 2010.

Levi, Isaac. 1967. *Gambling with truth*. A. Knopf.

Levi, Isaac. 1974. On indeterminate probabilities. *Journal of Philosophy* 71: 391–418.

Levi, Isaac. 1980. *The enterprise of knowledge: An essay on knowledge, credal probability, and chance*. The MIT Press.

Lewis, David. 1980. A subjectivist's guide to objective chance. In Richard C. Jeffrey (ed.), *Studies in inductive logic and probability*, vol. II, University of California Press.

Lewis, David. 1986. *On the plurality of worlds*. Blackwell Publishing: Malden, MA.

Maher, Patrick. 1993. *Betting on theories*. Cambridge University Press.

Maher, Patrick. 2002. Joyce's argument for probabilism. *Philosophy of Science* 69: 73–81.

Makinson, D. C. 1965. The paradox of the preface. *Analysis* 25: 205–7.

Moss, Sarah. 2011. Scoring rules and epistemic compromise. *Mind* 120(480): 1053–1069.

Moss, Sarah. forthcoming. Credal dilemmas. *Noûs* .

Nelkin, Dana K. 2000. The lottery paradox, knowledge, and rationality. *The Philosophical Review* 109(3): 373–408.

Neta, Ram. 2008. What evidence do you have? *British Journal for the Philosophy of Science* 59(1): 89–119.

Newell, Ben R. & Michael D. Lee. 2011. The right tool for the job? comparing an evidence accumulation and a naïve strategy selection model of decision making. *Journal of Behavioral Decision Making* 24(5): 456–481.

Pettigrew, Richard. 2012. Accuracy, chance, and the principal principle. *The Philosophical Review* 121(2): 241–275.

Pettigrew, Richard. 2013. A new epistemic utility argument for the principal principle. *Episteme* 10(1): 19–35.

Pettigrew, Richard. forthcoming. Accuracy, risk, and the principle of indifference. *Philosophy and Phenomenological Research* .

Pollock, John L. 1971. Perceptual knowledge. *The Philosophical Review* 80(2): 287–319.

Pollock, John L. 1974. *Knowledge and justification*. Princeton University Press.

Pollock, John L. 1995. *Cognitive carpentry*. Cambridge: MIT Press.

Popper, Karl. 1959. *The logic of scientific discovery*. London: Hutchinson & Co.

Pruss, Alexander R. 2012. Infinite lotteries, perfectly thin darts, and infinitesimals. *Thought* 1(2): 81–9.

Pruss, Alexander R. 2013. Probability, regularity, and cardinality. *Philosophy of Science* 80(2): 231–240.

Pryor, James. 2000. The skeptic and the dogmatist. *Noûs* 34(4): 517–549.

Pryor, James. 2005. There is immediate justification. In Matthias Steup & Ernest Sosa (eds.), *Contemporary debates in epistemology*, Blackwell.

Pryor, James. 2013. Problems for credulism. In Chris Tucker (ed.), *Seemings and justification: New essays on dogmatism and phenomenal conservatism*, Oxford University Press.

Rényi, Alfred. 1970. *Foundations of probability*. San Francisco: Holden-Day In.

Ross, Jacob & Mark Schroeder. 2014. Belief, credence, and pragmatic encroachment. *Philosophy and Phenomenological Research* 88(2): 259–288.

Roush, Sherrilyn. 2005. *Tracking truth: Knowledge, evidence, and science*. Oxford University Press.

Ryan, Sharon. 1996. The epistemic virtues of consistency. *Synthese* 109(22): 121–141.

Savage, Leonard J. 1971. Elicitation of personal probabilities and expectations. *Journal of American Statistical Association* 66(336): 783–801.

Seidenfeld, Theodore. 2004. A contrast between two decision rules for use with (convex) sets of probabilities: γ-maximin versus e-admissibility. *Synthese* 140(1-2): 69–88.

Shafer, Glenn. 1976. *A mathematical theory of evidence*. Princeton University Press.

Silins, Nicholas. 2008. Basic justification and the moorean response to the skeptic. In *Oxford studies in epistemology*, vol. 2, Oxford University Press.

Skyrms, Brian. 1980. *Causal necessity*. Yale University Press: New Haven.

Smith, Martin. 2010. A generalized lottery paradox for infinite probability spaces. *British Journal for the Philosophy of Science* 61(4): 821–831.

Steele, Katie. 2010. What are the minimal requirements of rational choice? arguments from the sequential setting. *Theory and Decision* 68(4): 463–487.

Sturgeon, Scott. 2008. Reason and the grain of belief. *Noûs* 42(1): 139–65.

Titelbaum, Michael. 2010. Tell me you love me: Bootstrapping, externalism, and no-lose epistemology. *Philosophical Studies* 149(1): 119–34.

van Fraassen, Bas. 1983. Calibration: A frequency justification for personal probability. In Robert S. Cohen & Larry Laudan (eds.), *Physics, philosophy, and psychoanalysis*, D. Reidel.

van Fraassen, Bas. 1984. Belief and the will. *The Journal of Philosophy* 81(5): 235–256.

van Fraassen, Bas. 1990. Figures in a probability landscape. In *Truth or consequences*, Kluwer.

Vineberg, Susan. 2014. Dutch book arguments. *Stanford Encyclopedia of Philosophy* Retrieved August 27, 2014: http://plato.stanford.edu/entries/dutch-book/.

Vogel, Jonathan. 2000. Reliabilism leveled. *Journal of Philosophy* 97(11): 602–23.

Vogel, Jonathan. 2008. Epistemic bootstrapping. *Journal of Philosophy* 105(9): 518–539.

Wagner, Carl. 2013. Is conditioning really incompatible with holism? *Journal of Philosophical Logic* 42(2): 409–414.

Weatherson, Brian. 2012. Knowledge, bets, and interests. In Jessica Brown & Mikkel Gerken (eds.), *Knowledge ascriptions*, 75–103. Oxford University Press.

Weintraub, Ruth. 2008. How probable is an infinite sequence of heads? a reply to williamson. *Analysis* 68(3): 247–250.

Weisberg, Jonathan. 2007. Conditionalization, reflection, and self-knowledge. *Philosophical Studies* 135(2): 179–197.

Weisberg, Jonathan. 2009. Commutativity or holism: A dilemma for conditionalizers. *British Journal for the Philosophy of Science* 60(4): 793–812.

Weisberg, Jonathan. 2010. Bootstrapping in general. *Philosophy and Phenomenological Research* 81(3): 525–48.

Weisberg, Jonathan. 2012. The bootstrapping problem. *Philosophy Compass* 7(9): 597–610.

Weisberg, Jonathan. 2013. Knowledge in action. *Philosophers' Imprint* 13(22): 1–23.

Weisberg, Jonathan. forthcoming. Updating, undermining, and independence. *British Journal for the Philosophy of Science* .

Wenmackers, Sylvia & Leon Horsten. 2013. Fair infinite lotteries. *Synthese* 190(1): 37–61.

White, Roger. 2005. Epistemic permissiveness. *Philosophical Perspectives* 19(1): 445–459.

White, Roger. 2006. Problems for dogmatism. *Philosophical Studies* 131(3): 525–557.

White, Roger. 2010. Evidential symmetry and mushy credence. In Tamar Szabo Gendler & John Hawthorne (eds.), *Oxford studies in epistemology*, vol. 3, 161–186. Oxford University Press.

Williamson, Timothy. 2000. *Knowledge and its limits*. Oxford University Press.

Williamson, Timothy. 2007. How probable is an infinite sequence of heads? *Analysis* 67(3): 173–80.