Epistemic Decision Theory's Reckoning[†]

CONOR MAYO-WILSON AND GREGORY WHEELER

Epistemic decision theory (EDT) employs the mathematical tools of rational choice theory to justify epistemic norms, including probabilism, conditionalization, and the Principal Principle, among others. Practitioners of EDT endorse two theses: (1) epistemic value is distinct from subjective preference, and (2) belief and epistemic value can be numerically quantified. We argue the first thesis, which we call *epistemic puritanism*, undermines the second.

§

Epistemic decision theory (EDT) is a reform movement within Bayesian epistemology that refashions rational choice theory into a wide-ranging defense of epistemic norms, including probabilism, conditionalization, and the Principal Principle.[1] Disciples of EDT endorse two theses: *epistemic puritanism*, which asserts that epistemic value should be absolved of subjective preferences, and *numerical quantifiability*, which maintains that epistemic value is numerically quantifiable. The problem, so we shall argue, is that puritanism puts the kibosh on quantifiability. For the soundness of the representation theorems that EDT must appropriate from traditional decision theory depend on features of a rational agent's subjective preferences. Take away those subjective features, as puritanism demands, and EDT is left without the means to establish a numerical representation of epistemic value.

1 The Origins of Epistemic Puritanism

It is one thing to say that your belief in p is stronger than your belief in q and quite another to say that your partial beliefs have precise numerical values. *Probabilism* is the even stronger thesis that a rational agent's synchronic partial beliefs are representable by a probability function.

The original argument for probabilism, given independently by Ramsey (1926) and, in greater detail, by de Finetti (1937), is that if an agent's partial beliefs are identified with the prices at which she is willing to buy and sell a finite number

[†]This paper was written in 2014 and presented to audiences at Bristol, Columbia, and the MCMP in Munich that year. This version of the paper, archived in August 2019, is longer than the original, for it includes the best of our failed attempts to appease journal referees over the years along with a few turns of phrase that pleased us. But our argument remains the same. In preparing this archived version, we have attempted to add references that have appeared since 2014 that directly fit into the paper, but have decided against refashioning the paper to reflect the hindsight that 5-plus years provides.

[1]Contemporary advocates include Jim Joyce (1998, 2005); Richard Pettigrew (2012, 2012, 2013b, 2016b) and Hannes Leitgeb with Pettigrew (2010); Hillary Greaves and David Wallace (2006); Jason Konek and Ben Levinsten (2016); Marcello D'Agostino and Corrado Sinigaglia (2010), although epistemic utility is not a new notion (Levi 1963; Levi 1967).

of gambles, then she exposes herself to the possibility of sure loss if and only if her partial beliefs fail to satisfy the axioms of (finitely additive) probability. The customary complaint about the Ramsey-de Finetti "Dutch Book" argument is that it relies on the prudential criterion of sure-loss avoidance. Since it is unclear that gambling losses are symptomatic of epistemic irrationality, this criterion may set the bar for rationality too high. Conversely, avoiding sure loss may arguably set the bar too low, as anyone with probabilistic beliefs is deemed rational regardless of the available evidence and truth.

De Finetti later developed a second criterion for rational belief, one that assesses the accuracy of an agent's elicited credences by proper scoring rules (de Finetti 1974). According to this scheme, an agent announces a number to express how strongly she believes a proposition p, and she does so on the understanding that she will be penalized by how far her announced "forecast" diverges from the truth-value, zero or one, of p. Such an agent has an incentive to report a number that accurately represents her credence in p, which is the *expected* truth-value of p. Here then is a purportedly epistemic criterion—accuracy—for rational belief. De Finetti further showed that these two methods for eliciting beliefs, the traditional Dutch Book argument and the Accurate Forecasting argument, are equivalent in the sense that partial beliefs that avoid sure loss are undominated by rival forecasts and vice versa.

Nonetheless, de Finetti's accuracy condition leaves many epistemologists unsatisfied.[2] Loss functions specify penalties that an agent pays in a particular currency, and de Finetti's theorem relies on a specific scoring rule, namely squared-error loss. But it is neither clear why epistemic rationality requires minimizing loss in some currency nor why one should minimize squared-error loss rather than loss measured by some other function.[3] An alternative account is needed.

This is the motivation for securing what Jim Joyce calls a "non-pragmatic" justification for probabilism. Joyce argues that any epistemic loss function, which he calls a measure of *gradational inaccuracy*, ought to satisfy certain axioms, and he goes on to show that an agent's beliefs are undominated with respect to such a measure of epistemic loss only if they satisfy the probability axioms. Joyce stresses that constraints on measures of gradational accuracy are justified by objective, purely *epistemic* arguments, in that they do not depend on any particular agent's preferences or subjective interests.[4] The recent rise in epistemic puritanism stems from

[2]See (Joyce 1998). Another case, albeit incomplete, for non-pragmatic foundations for probabilism is given by Roger Rosenkrantz (1981).

[3]Recent work has extended de Finetti's representation theorem for coherent forecasts beyond Brier's score. See (Predd, Seringer, Lieb, Osherson, Poor, and Kulkami 2009) for an extension to a large class of strictly proper scoring rules, and also (Schervish, Seidenfeld, and Kadane 2014) for an extension to general random variables.

[4]See (Joyce 1998, p. 586). As a second example, Easwaran claims, "[Joyce] is able to justify the probability axioms as constraints on rational credence *without assuming any constraint on rational preference* beyond dominance" (Easwaran 2015, p. 15, our emphasis). Similar claims that epistemic reasons cannot be grounded in subjective interest or preference are common throughout epistemology. For instance, Tyler Burge claims, "Reason has a function in providing guidance to

this non-pragmatic or *accuracy-first* defense of probabilism.

Underlying all of EDT, however, is the assumption that epistemic loss, or epistemic value, is numerically quantifiable. While epistemic decision theorists are keen to discuss which constraints an epistemic loss function should abide by—whether it should be continuous, convex, and the like—few defend the assumption that epistemic loss is a genuine quantity, and the arguments offered so far are unconvincing.

Leitgeb and Pettigrew, for instance, maintain that the distance between belief and the truth is quantifiable because both can be represented on the same scale.

> *Since truth and falsity [are] represented by real numbers, too, degrees of belief and truth values are comparable—they occupy the same quantitative or geometrical scale.*[5]

This argument is dubious, however, as two quantities can be represented within a single numerical scale without being comparable. By an appropriate choice of units, both wavelengths of light in the visible spectrum and audible frequencies of sound are representable by numbers between zero and one. Yet that hardly establishes there is a meaningful sense in which the color red is closer to the pitch Middle C than to the color blue. Isomorphic mathematical structures arise everywhere. Meaningful comparisons do not.

Joyce, who concedes that epistemic value may not be a precise quantity, nevertheless argues that it is useful to assume quantifiability:

> *I will speak as if gradational accuracy can be precisely quantified. This may be unrealistic since the concept of accuracy for partial beliefs may simply be too* ***vague*** *to admit of sharp numerical quantification. Even if this is so, however, it is still useful to pretend that it can be so characterized since this lets us take a "supervaluationist" approach to its vagueness. The supervaluationist idea is that one can understand a vague concept by looking at all the ways in which it can be made precise, and treating facts about the properties that all its "precisifications" have as facts about the concept itself. In this context a "precisification" is a real function that assigns a definite inaccuracy score $I(b,\omega)$ to each set of degrees of belief b and world ω.*[6]

The problem with this argument is that there are many ways of making inaccuracy "precise" that do not require using real numbers; degrees of inaccuracy might be represented by members of any ordered set, for example. To argue that any precisification of accuracy is numerical is to put cart before horse.[7]

truth, in presenting and promoting truth without regard to individual interest. This is why epistemic reasons are not relativized to a person or to a desire" (Burge 1993, p. 475).

[5] Leitgeb and Pettigrew (2010, pp. 211–212).

[6] Joyce (1998, p. 590).

[7] An additional challenge for Joyce is to reconcile his dual commitments to set-based Bayesianism and accuracy-based probabilism (Mayo-Wilson and Wheeler 2016).

Nevertheless, suppose for a moment that every precisification of accuracy were numerical. It still does not follow that epistemic value would be quantifiable. For instance, two people might both prefer oranges to apples but do so for very different reasons. One may prefer the taste of oranges to apples, while the other prefers the greater health benefits from eating oranges to eating apples. That oranges are a better fruit on a number of subjective dimensions does not establish there is an objective quantity, "*distance from the best fruit*", to compare apples and oranges by. Similarly, a precisification argument would not reveal anything about an objective epistemic quantity called "accuracy" that is independent of any agent's preferences.

Epistemic decision theorists might respond that it is simply a conceptual truth that accuracy is numerical. After all, measures of accuracy are intended to represent "distance from the truth," and various types of distance are fruitfully represented by real numbers.

Expressions like "their views are miles apart" and "he has a long way to go," however, are ubiquitous in natural language, and such metaphors provide no motivation for representing "distance" between opinions, guesses, or steps of a plan using real numbers. More still, it makes little sense to say one person's view is two-and-a-half times closer to the truth than another's. So even if accuracy is defined as "distance to the truth," it is not an obvious conceptual truth that accuracy is numerical. An argument is necessary.

One might concede the arguments for quantifying pure epistemic value are bad but still maintain that quantifiability is an innocent assumption. Surely, with a little effort, the standard representation theorems from decision theory could be amended to show that epistemic loss is numerical. We think not, and we shall argue that EDT founders on the irreconcilable demands of epistemic puritanism and numerical quantifiability. Although a modified EDT with non-puritanical foundations might be possible, substantial assumptions would need to be made about how credence and accuracy are measured along with some serious mathematical work. A simple repackaging of traditional decision theory that swaps "loss" for "accuracy" will not do.

Our central argument, presented in Section 3, is that comparisons of epistemic value fail to satisfy two different groups of conditions that are necessary for canonical representation theorems of value. The first group consists of a pair of assumptions that any numerical representation theorem of epistemic value must satisfy: totality and transitivity. The second group includes conditions that are required for EDT's devotion to proper scoring rules: continuity and independence. To set up our central argument, we review von Neumann and Morgenstern's (1944) canonical representation theorem in Section 2, and we return in Section 4 to make clear why EDT's commitment to making comparisons of expected epistemic value depends on the four axioms we target. In Section 5 we consider what would be involved in formulating an alternative representation theorem that avoided all of the axioms we target. Finally, we close in Section 6 with an overall assessment of EDT.

2 Representation Theorems for Numerical Utility

Why is numerical utility important to EDT? Isn't it possible to forgo numerical utility altogether and justify probabilism by cobbling together a list of intuitively compelling axioms for qualitative judgments of uncertainty? This was de Finetti's aspiration, who, in the course of proving a probability representation theorem from qualitative axioms and a cardinal utility function (de Finetti 1931), asked whether his qualitative postulates alone might entail the probability axioms. They do not (Kraft, Pratt, and Seidenberg 1959), and three options for augmenting de Finetti's qualitative axioms emerged: Savage's (1954), Anscombe and Aumann's (1963), and de Finetti's (1937). All of EDT's existing mathematical arguments use de Finetti's framework, however, and neither of the other two options are viable alternatives.[8] At the heart of de Finetti's framework is the assumption that beliefs can be scored in terms of cardinal utility.

Although de Finetti justified his use of cardinal utility by identifying monetary value and utility "within the limits of 'everyday affairs'" (de Finetti 1974, p. 82), viewing the utility of money as linear in monetary value is not a requirement of rationality.[9] Rather, what is required of a theory within the de Finetti tradition is a method for constructing an utility scale where numerical utilities on that scale have a clear-cut interpretation. The problem is that EDT uses de Finetti's framework without accounting for how to construct a meaningful scale for epistemic utility.

How might EDT construct such a scale? For traditional theories of rational choice the steps for constructing an effective scale for comparative judgments of value (preferences) are spelled out by von Neumann and Morgenstern's (1944) theory of utility: (1) identify preferences with choices, (2) postulate rationality axioms governing choices, and (3) prove a representation theorem showing that, when an agent's choice behavior satisfies those rationality postulates, she can be modeled *as if* she attached numerical utilities to the different options. We review these three steps in order.

[8]The first option, Savage's, extends de Finetti's qualitative axioms by enriching the state space. But Savage assumes that the state space is infinite and that an agent's probability judgments are derived from her subjective preferences. EDT rejects both assumptions, and adopting this line would require abandoning all existing mathematical arguments of EDT. The second option, Anscombe and Aumann's, enriches the outcome space from basic outcomes to von Neumann and Morgenstern lotteries. Unlike Savage, Anscombe and Aumann use roulette lotteries as an objective public standard to measure individual credences. But their approach is like Savage's in deriving numerical probability from subjective preferences. That leaves de Finetti's strategy, which effectively enriches the set of acts an agent considers—called *gambles*—to yield comparable real-valued rewards.

[9]For instance, *probability currency* (Smith 1961) is a common method for constructing a cardinal utility scale in which the value of rewards have an operationalizable interpretation. However, even the traditional coherence scheme admits a formal notion of constrained coherence (Vicig 2016) to mitigate against non-linear distortions introduced by varying the size of a stake.

2.1 Identifying Preference with Observable Choices

Within traditional decision theory an agent's (*i*) observable behavior (*choices*) are identified with (*ii*) her unobservable comparative value judgments (*preferences*) by way of (*iii*) an operationalizable *protocol* that explains which preferences are identified with which choices. Although authors in this tradition place different weight on these three components,[10] a behavioristic interpretation which identifies choice and preference remains commonplace.

According to one protocol, an agent is said to prefer an option x to another option y precisely if she chooses x in situations in which x and y are the only two available options and exactly one of x and y must be chosen.[11] If she would also choose y in these situations, then the agent is said to be indifferent between x and y. Otherwise, her preference of x to y is *strict*. Here x and y might be objects, like apples and oranges; they might be actions, like buying an apple or eating an orange; or they might even be states of affairs, such as whether there are apples in the bowl or oranges on the tree. For now, all that matters is that an individual can either indicate which of two options she prefers or be indifferent between the two.

There are two upshots of identifying preferences with observable choices. First, learning an individual's preferences requires observing her behavior, not her internal mental states. Second, for any two options x and y, an individual has a well-defined strict preference between x and y or is indifferent between the two. For either one chooses x when y is available or she does not. In the former case the agent prefers x to y, by the definition of preference, and in the latter she strictly prefers y to x.

We can put this succinctly by introducing some notation. Let $x \preceq y$ abbreviate the claim that an individual—say *you*—prefers y to x or is indifferent between the two. The above argument shows that, if preference is identified with your choices, then for any options x and y, either $x \preceq y$ or $y \preceq x$. This property is called *totality*.[12]

One may object that in practice you are not forced to make choices and that not all of your choices can be observed. Indeed, some decision theorists do not assume totality but rely instead upon weaker representation theorems, which with the other axioms below, entail that your options are assigned a range of numerical utilities rather than a single number. We postpone discussion of these weaker theorems

[10]For instance, Von Neumann and Morgenstern identify preference with a subjective feeling, available via introspection (von Neumann and Morgenstern 1944, §3.3.2), a position Ramsey (1926) among others flatly rejects; Sen (1982) views choice behavior to reveal rather than constitute preference; while others, particularly in economics, view the protocol to apply broadly enough to simply identify choice and preference. See for example (Gilboa 2009).

[11]This definition of preference is more restrictive than what is used in more recent research on "choice functions," as systematized by Amartya Sen (1971). In this more general setting binary preferences are derived from decisions in which more than two options are available. We restrict our attention to choices among two options for simplicity of exposition only.

[12]Authors who maintain that forced choices "reveal" but do not constitute preferences cannot justify totality unless one adopts the non-empirical assumption that when "a person chooses x rather than y, it is presumed that he regards x to be at least as good as y, and not that maybe he has no clue about what to choose and has chosen x because he had to choose something" (Sen 1973).

until Section 2.3, and return again in Section 3.1 to discuss their incompatability with EDT.

2.2 Axioms of Rational Choice

The totality of $\preceq$ follows from the definition of preference, but decision theorists typically argue that an individual's preferences ought to satisfy additional rationality postulates. We will focus on three: transitivity, continuity, and independence.

Most agree that preference ought to be *transitive*: if you prefer x to y and you prefer y to z, then you ought to prefer x to z. Transitivity is often justified by "money pump" arguments, which make two assumptions: if you strictly prefer x to y, then you would be willing to pay a small price to exchange y for x; and if you are unwilling to pay anything for a trade, then you are indifferent between the two. It is easy to show that if you have intransitive preferences and these two assumptions hold, then a clever trader can cause you financial ruin through a package of trades that you consider to be fair. Hence, rational preferences are transitive.[13]

Suppose now we wished to determine how much more you value an orange to an apple. To meaningfully quantify the degree to which you prefer an orange to an apple, the options of choice presented to you must be expanded to include not just apples and oranges but options like the following:

(L1) A die will be rolled. If a one is observed, you will be given an apple. Otherwise, you will receive an orange.

(L2) A fair coin will be tossed. You will get an apple if the coin lands heads, and an orange otherwise.

(L3) A die will be rolled. If a six is observed, you will receive nothing. Otherwise, you will receive an orange.

These three options are called *lotteries*. Lotteries are options in which different objects are awarded to an agent with different "objective" probabilities, which are determined by chance devices like coins, dice, and roulette wheels. The reason lotteries are important is because they allow an observer to determine how much more you prefer one option to another. If you prefer oranges to apples, then you will also prefer the first lottery to the second, as the first gives you a higher probability of getting an orange. However, depending upon how much you prefer oranges to apples, you might not prefer the third lottery to the second. In the third lottery there is some chance you get nothing. If you enjoy oranges much more than apples, getting nothing might be a risk you are willing to take. On the other hand, if you only slightly prefer oranges to apples, that chance might be unacceptable. Von Neumann and Morgenstern's idea was that by appropriately changing the probabilities

[13]The axioms of rational choice impose synchronic consistency constraints on preference, but there is a compelling case for rational violations of transitivity of preferences over time (Wheeler 2018, §1.2).

with which you are awarded different basic options, such as apples and oranges, an experimenter can determine how much you prefer one to the other.

Equipped with the concept of a lottery, we can now state the next rationality axiom: *continuity*. Consider two options x and y and some probability α between zero and one. Let $\alpha x \oplus (1-\alpha)y$ denote the lottery in which you win x with probability α and win y with probability $1-\alpha$. Then the continuity axiom says that if you strictly prefer y to x and you strictly prefer z to y, then there are probabilities α_1 and α_2 between zero and one such that (*i*) you strictly prefer y to the lottery $\alpha_1 x \oplus (1-\alpha_1)z$, and (*ii*) you strictly prefer the lottery $\alpha_2 x \oplus (1-\alpha_2)z$ to y.

What the continuity axiom rules out is that preference can be "infinitely" strong. If you think apples are awful but oranges are outstanding, then you might not prefer any lottery in which there is a chance, however small, of ending up with only an apple. If so, your preferences violate the continuity axiom.

However, if you are unwilling to take a gamble involving apples, then surely you ought not take the same gamble if the apple prize is replaced by your own death. Yet we routinely make choices that raise our risk of death—we drive cars and walk across busy streets—often in exchange for a small reward like an orange. Yet if death is only so bad, how bad could an apple be? On closer inspection, seemingly infinite losses (and gains) are really not infinite at all. Such is the empirical argument given for the continuity axiom.[14]

The last rationality postulate is *independence*.[15] Suppose you prefer oranges to apples and you are offered a choice between two lotteries. In the first, a fair coin is flipped and you get an orange if it lands heads. Otherwise, you receive a banana. The second lottery is exactly like the first except that you receive an apple if the coin lands heads. Because you prefer oranges to apples, intuitively you ought to prefer the first lottery to the second. This is what the independence axiom says. In symbols, if you prefer y to x, then for any probability α less than one and any third option z, you should also prefer the lottery $\alpha y \oplus (1-\alpha)z$ to the lottery $\alpha x \oplus (1-\alpha)z$.

With these four necessary conditions in hand—totality, transitivity, continuity, and independence—we now turn to the last step of the canonical argument that utility is numerical: proving a representation theorem.

2.3 Representation Theorem

Von Neumann and Morgenstern prove that if an agent's preference relation $\preceq$ satisfies the above core axioms (and technical fusses) then every simple option a can

[14] We are not committed to the continuity axiom, and in fact we think there are compelling arguments that rational agents *must* occasionally violate it. See Arthur Paul Pedersen (2014). Nevertheless, we discuss the axiom because EDT needs to provide a purely epistemic argument for it or an equivalent condition.

[15] The remaining axioms necessary for von Neumann and Morgenstern's result are technical fusses, including a reduction axiom for compound lotteries (i.e., lotteries of lotteries are lotteries), axioms about sure events (you should be indifferent to an apple and the lottery with the sure chance of winning an apple) and null events. For brevity, we omit these.

be assigned a numerical utility $U(a)$, and moreover, for any options b and c

$$b \preceq c \text{ if and only if } \mathbb{E}_U[b] \leq \mathbb{E}_U[c],$$

where $\mathbb{E}_U[b]$ is the *expected utility* of the option b. For any option b that is not a lottery, its expected utility is defined to be $U(b)$, i.e., its utility simpliciter. The expected utility of a lottery $\langle \alpha, b, c \rangle$ is then defined (recursively) to be $\alpha\mathbb{E}_U[b] + (1-\alpha)\mathbb{E}_U[c]$. Von Neumann and Morgenstern further show this numerical function is *unique* up to a linear transformation, which means that if U_1 and U_2 are two different ways of assigning numerical utilities to options, then there exist real numbers r and s such that $U_1(a) = r \cdot U_2(a) + s$ for every option a.

This uniqueness property is important and requires some explanation. Suppose we wish to quantify an agent's preferences among a finite number of basic options; no lotteries are considered. Then, assuming only that the agent's preferences are total and transitive, there is a trivial way to assign utilities: assign the least preferred object zero utility, the next utility one, the next two, and so on. However, there are many other ways of assigning numerical utilities: assign the least object utility two, the next utility four, the third utility eight, and so on. In fact, any increasing function will do. What we have described is a numerical utility function that preserves only so-called *ordinal* preferences.

Von Neumann and Morgenstern's theorem shows that the set of numerical utility functions that represent an agent's preferences is far narrower if one also assumes the continuity axiom. Why care? Suppose an experimenter has elicited your preferences among fruit and wants to predict whether you will go to a supermarket or to a local bodega this afternoon. Here, your choice depends not only upon your preferences among fruits, but also upon your judgments about the likelihood of finding those fruits in the two stores. Finally, suppose you maximize your *expected* utility, and that the experimenter knows your judgments about the likelihood of finding various fruits in the two stores. If the experimenter knows your utility function up to an affine transformation, then she can determine which store you will choose. In contrast, if only your ordinal preferences are known, then your choice cannot be determined.

This uniqueness is important for normative as well as for predictive purposes. If the experimenter wants to recommend a store to you and if her recommendation is for you to maximize subjective expected utility, it suffices for her to know your utility function up to linear transformation. Conversely, she might be unable to make a recommendation without such knowledge, as two utilities not related by a linear transformation might disagree about which options have higher expected value. This point is important for the remainder because epistemic decision theorists typically require that epistemic utility is measured by a proper scoring rule, and the definition of a proper scoring rule involves comparisons of expected utility.

Von Neumann and Morgenstern's theorem can be generalized in several ways, but perhaps the most important generalizations drop the totality assumption.[16] Drop-

[16] See R. Duncan Luce (1956) and Peter Fishburn (1991) for arguments to drop transitivity.

ping totality allows one to avoid *identifying* preference with behavior. But decision theorists must assume there is some relationship between preference and choice,[17] as a lack of procedural constraints on measuring preference translates into a lack of justification for the axioms of representations theorems.[18] Crude behaviorism is implausible, but on pains of abandoning decision theory's mathematical framework entirely, one cannot reject the fundamental pragmatist insight: how preference is measured determines the axioms we may justifiably assume it satisfies (or should satisfy).

Importantly, theorems that drop totality do not yield a unique numerical representation of preference, and they require other technical assumptions that are inimical to EDT. For example, Aumann assumes that the options of choice form a finite-dimensional vector space; Fishburn assumes that the set of equivalence classes of options modulo indifference is countable or that it contains a dense subset; similar results are obtained by Dubra, Maccheroni and Ok, and also by Seidenfeld, Schervish and Kadane.[19] These technical assumptions are not substantial in the original case concerning preferences, in which there are typically only finitely many basic options to choose. But they are substantial in the epistemic case, as we will see.

So much for the numerical representation of preference. We now turn to the question of whether epistemic value is likewise numerical.

3 Epistemic Decision Theory's Number Problem

Now what is the point of this numerical comparison? How is the number used?
– F. P. Ramsey (1929, p. 95)

EDT claims that every belief state can be assigned a numerical epistemic utility. In particular, many advocates of EDT identify "utility" with "accuracy", where accuracy is understood as the "distance" between one's belief and the truth. Can epistemic puritans adapt the arguments used in traditional decision theory to justify numerical measures of epistemic utility in general or accuracy in particular? We argue not. By Von Neumann and Morgenstern's theorem, if epistemic utility were numerical, then comparisons of epistemic value must satisfy all four postulates discussed in the previous section. We consider the axioms in order and argue that puritans can justify none of them.

[17] See chapter one of Walley (1991) for a perspicuous discussion.

[18] Before the late twentieth century, the only way to measure preference was via choice behavior, but neuroeconomics now proposes many alternative methods, including fMRI scans, eye-tracking studies, response-time analysis, and more. These alternative procedures for measuring preference could be used to justify numerical representations theorems, but then quantifiability of preference would be a descriptive fact (or idealization) rather than a normative thesis about avoiding incoherent behavior. Moreover, neuroeconomists' alternative measures of preference are often justified by showing they are correlated with measurements of choice behavior.

[19] Robert Aumann (1962); Peter Fishburn (1979, p. 18 and pp. 29–30); Juan Dubra, Fabio Maccheroni, and Efe A. Ok (2004); and Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane (1995).

3.1 Totality

First, there is totality. Generally, EDT presumes that an epistemic utility function takes as arguments pairs of the form $\langle b, \omega \rangle$, where b is a belief and ω is a truth-value assignment.[20] The intended interpretation of $\langle b, \omega_1 \rangle \prec \langle c, \omega_2 \rangle$ is "b is strictly more accurate in world ω_1 than c is in ω_2." We consider the special case of when b and c are compared in the same world ω, in which case we write $b \prec_\omega c$ if b is strictly more accurate than c in ω, and write $b \preceq_\omega c$ if b is at least as accurate as c in ω. For the moment, we remain agnostic about what a belief state is and whether beliefs are best represented by sets of propositions, probability functions, qualitative likelihood orderings, or some other type of mathematical object. The point of this setup is to deal EDT the strongest hand it can play.

The epistemic analog of totality of preference is that, for any two belief states b and c and any state of the world ω, either $b \preceq_\omega c$ or $c \preceq_\omega b$ (or both). What could justify this assumption? Puritans cannot resort to the decision theorist's argument that totality follows from the definition of preference, for there is no epistemic analog of a forced-choice scenario. Consider for example those who identify epistemic utility with accuracy. Accuracy is intended to represent the distance between b and the truth, and this distance is supposed to be a fact about the world, not about someone's preferences. It does not matter whether an agent would choose to have beliefs b rather than c if she knew the truth ω. So the pragmatic argument for totality does not extend naturally to the epistemic domain.

Is there any argument that comparisons of accuracy are total? Take the "distance to the truth" metaphor seriously for a moment and consider the following question: how far is Berlin from Paris and London? Even though this question is about distance, most readers would be tempted to answer not with a single number but with two: one to quantify the distance (in some specified unit of length) between Berlin and Paris, and another for the distance between Berlin and London. You may have reasons to collapse the pair of numbers into a single one—if you were a travelling salesperson tallying your miles, say—but there is no canonical way of doing so. How two numerical distances are combined into a single number depends upon ones interests and values.

An analogous situation arises in epistemology. Consider two propositions, p and q. Suppose p is the proposition "Most elements are metals" and q is the proposition "Taft wore a handle-bar mustache." Suppose Alison is skeptical of p but strongly believes q. How close are Alison's beliefs to the truth? A reasonable answer is, "Alison's beliefs about basic chemistry are off the mark, but her beliefs about the 27th President of the United State's facial hair are surprisingly accurate." Just as with distance, quantifying the accuracy of Alison's views will require specifying the relative importance of having accurate beliefs about chemistry to having accurate beliefs about Taft. This specification is easy for a pragmatist: one's prefer-

[20]Alternatively, Pettigrew conceives the pair to consist of your credal function, b, and the ideal credal function of an oracle that assigns 1 to a truth and 0 to a falsehood. Our argument does not hang on this distinction.

ences determine the relative importance of different propositions. But for puritans, there is no purely epistemic grounds for attaching weights to different propositions and averaging their accuracy.

Here is why. Epistemic goods, traditionally understood, are few in number; they include acquisition of true beliefs, avoidance of error, and justification. Further, epistemic decision theorists claim that credences can be accurate but unjustified and vice versa.[21] Thus, to show that numerical measures of accuracy are non-epistemic, it suffices to show that they encode values other than acquisition of truth and avoidance of error. As we have shown, numerical measures of accuracy quantify how a gain in accuracy with respect to one proposition is to be weighed against loss in accuracy with respect to another. Thus, numerical measures of accuracy, even those that weight all propositions equally, implicitly quantify the relative importance of the *contents* of different propositions, which is a value judgment about something other than proximity to truth.

What does this have to do with totality? Suppose Bill strongly believes "Most elements are metals" but doubts that "Taft wore a handle-bar mustache." If epistemic goods do not determine which proposition is more important or determine that they are equally important, then Alison's and Bill's beliefs are *incomparable* in terms of accuracy: Alison's views about Taft are more accurate than Bill's, but her beliefs about chemistry are less so. Thus, epistemologists who maintain that accuracy is the sole epistemic aim cannot defend the claim that comparisons of accuracy are total unless there is only one proposition under consideration.[22]

Puritans who wish to employ von Neumann and Morgenstern's theorem, therefore, must expand the list of epistemic goods to defend the claim that the relative importance of *all* propositions can be compared. Candidate epistemic goods include evidential support, explanatory breadth, fruitfulness, and other theoretical virtues discussed in the philosophy of science. But there are two reasons to doubt this strategy will yield a justification of the totality axiom for comparisons of epistemic utility. Neither of these arguments is new, but they are worth repeating.

First, new goods create new incomparabilities. Suppose it is more important to have accurate beliefs about p than q if p has greater explanatory breadth than q. Then to argue that comparisons of epistemic utility are total, one must show that considerations of explanatory breadth eliminate all incomparabilities. This is doubtful. Do true hypotheses of particle physics have more or less explanatory breadth than ones about spacetime? Both types of hypotheses explain a wide variety of phenomena, but of different sorts. In general, it is not meaningful to compare the evidential support, explanatory breadth, and other features of scientific hypotheses that are supported by widely different types of evidence and that explain completely different phenomena.

Second, different epistemic goals conflict. A proposition may have greater explanatory breadth than another but less evidential support. Even if different dimensions of epistemic value are always comparable, which is doubtful, one confronts

[21] See (Joyce 1998, pp. 591–592).

[22] See Pettigrew (2013a) for a defense of the view that accuracy is the sole epistemic good.

one of two dead-ends: either comparisons are exclusively ordinal and one faces Arrow-style impossibility results for combining different dimensions of epistemic value, or comparisons are "beyond ordinal" in which case the question of whether epistemic value is numerical is begged and answered.[23]

Epistemic decision theorists might try to employ one of the representation theorems that do not require totality. However, recall from Section 2.3 that these theorems require additional technical assumptions, like that the set of options is countable (modulo indifference), or forms a finite dimensional vector space, or contains a countable dense subset. The first assumption is false if belief states include—as most epistemic decision theorists assume—all uncountably many probability functions on an algebra. The most plausible defense of the second and third assumptions is to argue that belief states are numerical; in Section 3.4, we show such an argument is unavailable to puritans.

Representation theorems that drop totality also do not yield a unique utility function up to linear transformation, and that uniqueness is necessary to justify meaningful comparisons of *expected* utilities, as we shall return to argue in Section 4. One might object that EDT does not claim there is a unique epistemic utility function (even up to linear transformation), and in fact, proponents of EDT are quick to point out that their arguments work for any proper scoring rule. However, this response raises a question that goes to the heart of EDT: how ought we interpret the lack of uniqueness of an epistemic utility function? If accuracy is the sole epistemic good, yet the measurement of accuracy is not unique, what are the purely epistemic grounds for choosing one measure of accuracy over another? Pettigrew answers by endorsing *subjectivism*: "Each continuous and additive strictly proper inaccuracy measure is an acceptable measure for an agent to adopt as her own subjective inaccuracy measure" (Pettigrew 2016a, p. 75). Yet, if one concedes that accuracy is a function of an individual's subjective interests, the game is up; for then EDT collapses to traditional decision theory, where the contrast between "subjective inaccuracy" and "subjective preference" expresses a distinction without a difference.

Puritans, however, might retort that there is an important difference between *value* and *preference*. Even granting this response, it does not explain why value judgments must be total. To see why, consider Sousa's (1974) "small-improvement" criterion for incomparability.[24] Even if you prefer viewing Rembrandt's paintings of apples to listening to George Gershwin's *An American in Paris*, you are likely reluctant to say one experience is more valuable than the other. So either they are equal in value, or they are incomparable. To show that that they are not equal in value, consider a slight improvement to a given performance of *An American in Paris*, one where the kettledrums are ever so slightly better in tune. Suppose this slightly improved performance of Gershwin is more valuable to you than the original performance. But then if the first Gershwin performance were equal in value to a given viewing of Rembrandt's work, the improved performance of Gershwin

[23] See Samir Okasha (2011).

[24] See (Chang 2002) for a discussion and history of this argument.

would be more valuable than viewing the Rembrandt. That strikes many as unintuitive: many are reluctant to judge any sufficiently competent performance of Gershwin to be more valuable than a viewing of Rembrandt's work. So the original reluctance in comparing Gershwin and Rembrandt is better explained by incomparability of value, rather than equivalence.

Analogous examples work in the epistemic case. Let p be a true proposition of mechanics and q a true proposition of genetics. Suppose Alison's credences in p and q are representable by the numbers .4 and .6 respectively, whereas Bill's credences are exactly reversed. One might lack a firm judgment that Alison's credal state is more accurate than Bill's or vice versa. So Alison and Bill's credences might be equally accurate, or they might be incomparable. Now consider Carole whose credences in p and q are representable by the numbers .4 and .61 respectively. Carole's beliefs are more accurate than Alison's, but it's not clear they are more accurate than Bill's. After all, Bill's credence in p is still more accurate than Carole's. But if Carole's beliefs are not obviously more accurate than Bill's, then this provides good evidence that our original lack of judgment about the relative accuracy of Alison and Bill's credences was motivated by incomparability rather than equivalence.

The following simplification of Anderson's (1987) theory of values explains why some values are incomparable and behave in the way Sousa's criterion describes. Say one *values* some object, person, or state of affairs S if one (1) acts in particular ways to acquire, preserve, promote, or bring about S and (2) feels particular emotional responses upon acquiring and losing S. For example, we value our friendships because we act in certain ways to promote them (e.g., talking by phone, writing letters, etc.) and we would feel sad and nostalgic upon losing friends. We value our jobs differently: professors, for example, publish papers and teach classes. And we would feel anxious and fearful about the future upon loss of our jobs.

This simplification of Anderson's theory can explain why certain values seem comparable whereas others do not. Consider two apples, which are identical except that one is a bit sweeter than the other. Intuitively, the two apples are comparable, and Anderson's theory explains why. We would take similar actions to enjoy them, and our feelings upon acquiring or losing them would be similar. We might pay more for the sweeter apple, enjoy the taste of it more than the taste of the other apple, or feel more upset bruising it than the other.

In contrast, our friendships and our careers sometimes seem incomparable in value, and Anderson's theory again explains why. We take *qualitatively different* actions to promote our friendships than we do to promote our careers. We also feel differently upon losing friends than we do upon losing our jobs. These differences are not of intensity, but of kind. Of course, sometimes we *do* choose (and have preferences) among careers and friendships; we might move across the country to maintain our careers, thereby undermining our friendships. But that does not entail that we value our careers more than our friends, as choice is but one way that value is measured. Thus, while some might value a friendship more than a career or vice versa, it is not *rationally mandated* to have a totally-ordered value relation over careers and friends.

Let's return to EDT. Must the accuracy of two belief states be comparable? Not at all. Suppose Alison is a physicist. How might she value accurate belief about a mathematical conjecture versus that of an empirical hypothesis about particle physics? Consider the actions Alison might take to promote accurate beliefs in the two cases. In the mathematical case, Alison will prove theorems, find intermediate lemmas, and develop novel definitions. In the experimental case, Alison will test measuring devices, record data, and perform statistical analyses. Alison's actions in the two cases differ fundamentally, and so the value of accuracy in the two cases differs.

Puritans might object that Anderson's theory is not well-suited for comparing *epistemic* value, as emotions are irrelevant in epistemology. However, our argument focused exclusively on value-promoting actions, not emotions. So at this point, we think the burden is on EDT to develop a coherent theory of epistemic value and to explain why comparisons of such value are total. We have considered the most plausible arguments for totality, and none works.

To recap, we first observed that the decision theorist's traditional defense of totality of preference does not extend to matters of pure epistemic value. We then noted that, using a restricted set of epistemic goods (i.e., truth, justification, and error-avoidance), beliefs in different propositions may not be comparable. Then we argued that widening the set of candidate "epistemic" goods (e.g., fruitfulness, simplicity, etc.) only complicates the matter. After that, we argued that criteria often used to show two values are incomparable (e.g., Sousa's) speak against the totality of ordering of epistemic value, and finally we summarized a variant of Anderson's theory of value that explains when and why incomparability arises.

3.2 Transitivity

Consider transitivity next. If $a \preceq_\omega b$ and $b \preceq_\omega c$, does it follow that $a \preceq_\omega c$? Money-pump arguments used to justify transitivity of preference are inapplicable, as these take for granted pragmatic incentives that puritanism forgoes. So to investigate the plausibility of transitivity in EDT, consider again the case in which epistemic utility is identified with accuracy.

	p	q	r
Alison	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$
Bill	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
Carole	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{8}$

Pick three true propositions, p, q and r. Imagine that, for all that Alison, Bill, and Carole know, the three propositions might all be false, some true and some false, or all true. Suppose the probabilities that Alison, Bill, and Carole assign to the three respective propositions are as described in the table above. Whose beliefs are most accurate?

A natural thought is that one probabilistic belief state is more accurate than another if the former is closer to the truth-values of most of the propositions.[25] When comparisons are made in this fashion, Bill's beliefs are closer to the truth than Alison's. Alison believes p to degree $\frac{1}{2}$ whereas Bill believes it to degree $\frac{3}{4}$, and the truth-value of p is one. Moreover, Alison believes q to degree $\frac{1}{4}$, whereas Bill believes it to degree $\frac{1}{2}$; again, the truth-value of q is one. Since Bill's degrees of belief are closer on two of three dimensions, it follows that his beliefs are closer to the truth than hers according to this scheme. By similar reasoning, Carole's beliefs are closer to the truth than Bill's: Carole has more accurate beliefs than Bill with respect to q and r. The problem is that Alison's beliefs are closer to the truth than Carole's, as Alison has more accurate beliefs with respect to p and r. So, transitivity fails.[26]

One might counter that this example does not show that comparisons of inaccuracy are intransitive but instead shows that naive tallying procedures violate transitivity. We agree. However, in order to avoid violations of transitivity, one needs to specify how acquiring an accurate belief with respect to one proposition is balanced against inaccuracies elsewhere, and this balancing act must be done carefully to avoid "accuracy cycles" like the one above. Yet it is precisely these trade-offs that make little sense to puritans, since they explicitly deny the importance of an agent's subjective preferences about whether an accurate belief with respect to p is more or less important than the same for q.

One may continue to press, however, by proposing an alternative to tallying which avoids intransitivity. For each investigator, one might instead calculate the degrees to which her beliefs in p, q and r are inaccurate and sum the result. Calculating a sum does not require appealing to an agent's subjective preferences about which propositions are most important because an unweighted sum may be viewed as treating the propositions "equally." Although we think that treating propositions "equally" is an ambiguous requirement, let us grant for now that this proposal for combining degrees of inaccuracy is privileged. The problem is that doing so runs afoul of another axiom: independence.

3.3 Independence

Recall that the independence axiom for preferences requires that the space of options contain lotteries. What is the epistemic analog of a lottery? At the most abstract level, a lottery is a triple, $\langle \alpha, x, y \rangle$, where α is some number between zero and one, and x and y are options. In the traditional, pragmatic case, the number α represents a probability that determines how likely different prizes are awarded in the lottery.

[25] An analogous notion of closeness has been used in the case of full-belief by (Easwaran and Fitelson 2015).

[26] This type of example is familiar to decision and social choice theorists. The former sometimes use similar examples to argue that there are intuitively acceptable violations of transitivity when objects have several different dimensions of value. The latter use it to show how cycles may arise in social preference if pairwise majority rule is used to rank options.

What about the epistemic case? Here the options x and y are replaced by belief states b and c. Whatever α represents, the resulting epistemic lottery $\langle \alpha, b, c \rangle$ must be an object whose accuracy can be compared to that of beliefs, in the same way that preferences are defined among lotteries. If lotteries were belief states, then this would be straightforward. So one proposal is to identify an "epistemic lottery" involving b and c with a belief state that is "intermediate" between b and c.[27]

The concept of an intermediate belief state makes sense in some contexts. Suppose $b = .6$ and $c = .4$ are the probabilistic credences that two agents assign to the proposition that it will rain tomorrow. Then an intermediate belief state between b and c is some number between .4 and .6. In particular, one might identify the "epistemic lottery" $\langle \alpha, b, c \rangle$ with the belief state $\alpha b + (1-\alpha)c$. If α is close to one, then the intermediate state is close to b. If α is close to zero, then it is close to c. So one can assess the epistemic utility of such lotteries in the same way one assesses the utility of probabilistic credences.

Unfortunately, the suggestion to treat propositions "equally" by summing accuracy scores violates independence if an epistemic lottery is understood in this way.

	p	q
Alison	$\frac{3}{4}$	$\frac{1}{4}$
Bill	$\frac{1}{4}$	$\frac{1}{2}$
Carole	1	0

Suppose Alison, Bill, and Carole's credences are as indicated in the table to the left. Further, assume both p and q are true. Now consider the policy that propositions ought to be treated "equally" by summing degrees of inaccuracy. According to the epistemic decision theorist's favorite measure of inaccuracy, squared-error loss, Alison's beliefs are more accurate than Bill's. Alison's beliefs are inaccurate to degree $(1-\frac{3}{4})^2 + (1-\frac{1}{4})^2 = \frac{5}{8}$, whereas Bill's inaccuracy score is $(1-\frac{1}{4})^2 + (1-\frac{1}{2})^2 = \frac{13}{16}$, and $\frac{5}{8} < \frac{13}{16}$. Thus, if the independence axiom were true, any belief state that is "intermediate" between Alison's and Carole's would be more accurate than the corresponding intermediate state between Bill's and Carole's beliefs. But if $\alpha = \frac{1}{25}$, then the mixture of Alison's and Carole's beliefs is less accurate than the corresponding mixture of Bill's and Carole's.[28]

The epistemic decision theorist thus faces a trilemma. One way of treating propositions equally is by our tallying method, but this violates transitivity. Another way of treating propositions equally is to sum inaccuracy scores, but that violates independence. Lastly, there are strategies that assign variable weightings to propositions which would satisfy both transitivity and independence, but there is

[27] Of course, in the same way that a pragmatic lottery with fruit prizes need not be itself a fruit, an epistemic lottery need not be a belief state. We return to this point at the end of the next section.

[28] If $\alpha = \frac{1}{25}$, then the belief state intermediate between Alison's and Carole's beliefs has a credence of $(\frac{1}{25} \cdot \frac{3}{4}) + (\frac{24}{25} \cdot 1) = \frac{99}{100}$ in p and a credence of $(\frac{1}{25} \cdot \frac{1}{4}) + (\frac{24}{25} \cdot 0) = \frac{1}{100}$ in q. The inaccuracy of that belief state is $(1-\frac{99}{100})^2 + (1-\frac{1}{100})^2 = .9802$. In contrast, the state intermediate between Bill's and Carole's has a credence of $(\frac{1}{25} \cdot \frac{1}{4}) + (\frac{24}{25} \cdot 1) = \frac{97}{100}$ in p and a credence of $(\frac{1}{25} \cdot \frac{1}{2}) + (\frac{24}{25} \cdot 0) = \frac{1}{50}$ in q. The inaccuracy of that state is $(1-\frac{97}{100})^2 + (1-\frac{1}{50})^2 = .9613 < .9802$. Similar counterexamples can be generated for any strictly proper scoring rule, as propriety entails strict convexity.

no purely epistemic motivation for assigning such unequal weights to propositions.

3.4 Continuity

We now turn to the continuity axiom. A quick gloss of this axiom might provide the epistemic puritan with some confidence that its analog is satisfied in the epistemic case. Recall, in the pragmatic case, the continuity axiom says roughly that no option is infinitely preferable to another. In the epistemic case, the rough analog is that no belief state is infinitely more accurate than another. A puritan might be perfectly content with such an assumption, as intuitively the worst belief state is to be maximally confident in all the false propositions, and such an epistemic state, while regrettable, surely does not seem “infinitely” bad.

Unfortunately, this argument assumes that there is some state of “maximal” confidence. Accuracy-first arguments for probabilism, however, often assume that belief is real-valued (and hence, unbounded) and attempt to prove that a rational belief is bounded.[29] Furthermore, some have argued that, in order to avoid lottery-like paradoxes, rational belief states must be “tiered” in the sense that some beliefs must be “infinitely” stronger than others.[30] But if some belief state is infinitely stronger than others, degrees of inaccuracy might likewise be infinitely better or worse than others. Finally, if there are infinitely many propositions under consideration, the assumption that epistemic loss is always finite is less plausible,[31] and it is inconsistent with weighting the inaccuracy of propositions “evenly” in calculations of the total inaccuracy of a belief state.

There is, however, an even more substantial problem with justifying the continuity and independence axioms. The epistemic analog of these axioms requires an epistemic analog of a lottery. Above, we identified an epistemic lottery with an intermediate belief state, and to define an intermediate belief state, we assumed that beliefs were probabilistic. Unfortunately, the entire point of EDT is to justify epistemic norms, probabilism being chief among the lot. Thus, on pains of circularity, one cannot argue that epistemic utility is numerical by assuming that beliefs are probabilistic.

Of course, one might argue that the above justification of the existence of intermediate belief states works under the weaker assumption that beliefs are numerical; one can take weighted-averages among numbers, even if those numbers are not probabilities. But why assume that belief is numerical? Unlike Ramsey and de Finetti, epistemic puritans do not identify degrees of belief with betting odds or fair prices.

Puritans might try to argue that rational agents make qualitative comparisons of probability, such as “Rain tomorrow is more likely than snow.” Moreover, there are theorems showing that, as long as those qualitative, comparative judgments satisfy

[29] See (Joyce 1998, p. 598) and Richard Pettigrew (2016a, §4.2).

[30] See Horacio Arlo-Costa and Arthur Paul Pedersen (2012).

[31] This criticism was suggested to us in personal conversation by Jonathan Livengood. Mikayla Kelley has since looked at extending epistemic loss to infinite option sets (Kelley 2019)

certain rationality axioms, then degrees of belief will be numerical. But there are two problems with this proposal. First, those rationality axioms (e.g., that comparative judgments of probability are transitive) are typically justified by the same pragmatic arguments that puritans disavow (e.g., money pumps). So epistemic decision theorists must argue that the axioms of qualitative, comparative probability embody purely epistemic constraints.

More importantly, the most uncontroversial axioms of comparative probability (e.g., totality, transitivity, and qualitative additivity) do not guarantee that degrees of belief admit *unique* numerical representation. Thus, even if belief states b and c are numerically representable, one cannot define the epistemic lottery $\alpha b \oplus (1-\alpha)c$ to be *the* weighted average of the numbers representing b and c, as such an average may not be unique. Yet without uniqueness, comparing of the accuracy of epistemic lotteries is meaningless.

To see why, consider a simple case with one proposition under investigation, p, and let a, b, and c be three belief states such that p is: (1) less likely than its negation according to a; (2) equally likely as its negation according to b; and (3) more likely than its negation according to c.[32] Although all three belief states are numerically representable, there is no reason to prefer representing c by the probability function $c(p) = \frac{9}{15}$ or $c(p) = \frac{3}{4}$. So even if a were uniquely represented by the function $a(p) = \frac{1}{3}$—putting aside for the moment that it is not—it would be indeterminate whether the intermediate belief state $d = \frac{1}{2}a \oplus \frac{1}{2}c$ assigned greater credence to p than does b, as d can be represented by both the credence function $d(p) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{9}{15} < \frac{1}{2} = b(p)$ and by $d(p) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{3}{4} > \frac{1}{2} = b(p)$.

Why does this matter for EDT? Suppose p is true. Then, according to EDT, b is more accurate than d if and only if b assigns greater credence to p than does d. Thus, b might be more or less accurate than d, depending upon how c is represented. Because the uncontroversial axioms of qualitative probability do not guarantee a unique numerical representation, it follows that they do not allow one to compare the accuracy of b with that of d.

Can EDT justify axioms of qualitative probability that guarantee uniqueness?[33] None of the axioms we know have plausible epistemic defenses that would not simultaneously further undermine the motivation for EDT. Some axioms guaranteeing uniqueness entail the number of propositions under consideration is infinite, and this consequence is inconsistent with EDT's current mathematical framework. Other assumptions guaranteeing uniqueness require intricate constraints on both one's qualitative probability judgments and the richness of propositions under investigation; these axioms have no plausible epistemic defense.[34]

---

[32] Formally, a, b, and c are represented respectively by the three orderings $\precsim_a, \precsim_b$ and $\precsim_c$ on the set of propositions $\{\bot, p, \neg p, \top\}$ such that (1) $\bot \prec_a p \prec_a \neg p \prec_a \top$, (2) $\bot \prec_b p \sim_b \neg p \prec_b \top$, and (3) $\bot \prec_c \neg p \prec_c p \prec_c \top$.

[33] See (Fishburn 1986, p. 338-339).

[34] For instance, (Luce 1967) shows that uniqueness is guaranteed if for every quadruple of events A, B, C, and D such that $A \cap B = \emptyset$, $A \succ B$, and $B \succeq D$, there are events C', D' and E such that $E \sim A \cup B$, $C' \cap D' = \emptyset$, $C' \cup D' \subseteq E$, $C \sim C'$, and $D \sim D'$.

Finally, even if the standard axioms guaranteeing unique numerical representation of belief could be given a purely epistemic defense, they entail not only that degrees of belief are numerical; they entail probabilism. So epistemic decision theorists face another dilemma. If they accept only the uncontroversial axioms of qualitative probability, belief need not be uniquely numerically representable, and it is unclear how to define an "intermediate belief state" or "epistemic lottery." If they accept the controversial axioms of qualitative probability that guarantee uniqueness, then Joycean-style arguments for probabilism are redundant.

There is one last possible out for the puritan, namely, to deny that an epistemic lottery should be interpreted as an intermediate belief. But this proposal only adds to EDT's woes.

For consider the general case of comparing accuracy of belief/world pairs. How should one interpret epistemic lotteries $\alpha\langle b, \omega_1\rangle \oplus (1-\alpha)\langle c, \omega_2\rangle$ of such pairs? Imagine an experimenter can pick a green or red apple to show a subject, and suppose that, via an independent method of adjusting the ambient lighting, she can control the subject's beliefs about the proposition p, "The apple is red." Such an experimenter could offer her subject the choice of (*i*) having belief b in world ω to (*ii*) having belief c in world ω_1 with objective chance α and having belief d in world ω_2 with objective chance $1-\alpha$. In such limited and highly artificial experimental conditions, a subject could say which lottery she prefers, and her preferences over such lotteries would be well-defined and total because her preferences would be her choices.

For puritans, however, the subject's preferences are irrelevant: puritans must justify why option (*i*) can be said to be more or less objectively accurate than option (*ii*). But does the subject who always has credence $\frac{3}{5}$ that the apple is red have more or less accurate beliefs than the subject who has a fifty-fifty objective chance of being in a red-apple world with credence one in p and of being in a green-apple world with credence one in $\neg p$? One cannot, without begging the question, say that the accuracy of an epistemic lottery so-described is an average of the numerical inaccuracies of its components. If epistemic lotteries are not belief states, what justifies the claim that comparisons of inaccuracy are meaningful, let alone that comparisons satisfy totality, transitivity, continuity, and independence? It is EDT's burden to say how to interpret an epistemic lottery if not as an intermediate belief and to show that the proposed interpretation satisfies the axioms necessary for some representation theorem.

4 Proper Scoring Rules and Expected Epistemic Value

Given the multitude of representation theorems,[35] why focus on von Neumann and Morgernstern's theorem? Disciples of EDT generally assume not only that epistemic value is quantifiable, but that it is also measured by a *proper scoring*

[35] David H. Krantz, Duncan R. Luce, Patrick Suppes, and Amos Tversky (1971) prove representation theorems for all sorts of quantities, such as length, pitch, color, among others.

rule. However, proper scoring rules work in the manner that EDT requires only if comparisons of *expected* epistemic value are meaningful. In this section, we argue that, given the theoretical commitments of EDT, all of the available options to justify meaningful comparisons of expected epistemic value employ von Neumann and Morgenstern's framework.

What is a proper scoring rule? Proponents of EDT assume (i) that every belief-world pair $\langle b, \omega \rangle$ can be assigned a numerical degree of epistemic utility $U(b, \omega)$, and (ii) every belief state b assigns a numerical strength of belief $b(\omega)$ to every state of the world ω. So given two belief states b and c, we can calculate the expected epistemic value of c relative to b as follows:

$$\mathbb{E}_{U,b}(c) = \sum_{\omega \in \Omega} b(\omega) \cdot U(c, \omega)$$

Then the measure of epistemic value U is called *proper* if $\mathbb{E}_{U,b}(b) \geq \mathbb{E}_{U,b}(c)$ for any two distinct belief states b and c. A score is called *strictly proper* if the inequality in the last sentence is always strict. EDT devotees often claim that if epistemic value is measured by a strictly proper scoring rule, then from the perspective of someone who in belief state b, it looks uniquely rational to be in b, as all other beliefs will be judged to have strictly lower expected epistemic value.[36]

This last move, however, is too quick, for it assumes there is some relationship between epistemic rationality and *subjective expected* epistemic utility. In traditional decision theory, the close link between expected value and rational preference is justified by representation theorems. These theorems often contain von Neumann and Morgenstern's postulates, which is precisely why we have focused on von Neumann and Morgenstern's theorem. However, perhaps one can get away with weaker assumptions. To argue that proper scoring rules are relevant to epistemic rationality, puritans must endorse the following principle:

Bridge Principle. If an agent's beliefs and values are represented by the functions b and U, respectively, then she ought to value belief state c above d if and only if the expected epistemic value $\mathbb{E}_{b,U}(c)$ of c is greater than that of d.[37]

Without this principle, proper scoring rules seem irrelevant, as they are defined in terms of comparisons of expected epistemic value. Why endorse the bridge principle? Here's one thought: given a belief state b, we can think of any belief state c as if it were a lottery over various belief-world pairs. For instance, suppose there are two states of the world, $\omega_H =$ Heads and $\omega_T =$Tails. Let b be the credence function such that $b(\omega_H) = \frac{1}{3}$. The epistemic lottery associated with belief state c is $\frac{1}{3}\langle c, \omega_H \rangle \oplus \frac{2}{3}\langle c, \omega_T \rangle$. The bridge principle follows immediately from the conjunction of von Neumann and Morgenstern's theorem and two additional principles:

[36] See (Oddie 1997; Greaves and Wallace 2006; Gibbard 2007; Joyce 2009).

[37] (Pettigrew 2015) defends the Bridge Principle, which he calls EUC, in a different way than we consider here. Space prevents us from critiquing Pettigrew's argument in detail, but we think there is no purely epistemic argument that justifies the additivity assumptions of the theorem he employs.

A1. If an agent's beliefs and values are represented by the functions b and U, respectively, then she ought to value belief state c over d if and only if she values the epistemic lottery associated with c over the epistemic lottery associated with d.

A2. The agent's beliefs and values are represented by the functions b and U, respectively, and her judgments of epistemic value obey von Neumann and Morgernstern's axioms.

Although A1 is fairly intuitive, the Bridge Principle might still be false if A2 fails. This is why von Neumann and Morgenstern's theorem is so important to EDT. Suppose there are two epistemic lotteries L_1 and L_2 such that an agent values L_2 at least as much as L_1, or in symbols, $L_1 \preceq L_2$. However, if the agent's values violate the continuity or independence axiom, by von Neumann and Morgenstern's theorem, it is possible that the expected epistemic utility of L_2 is nonetheless *lower than* L_1, or in symbols, $\mathbb{E}_U(L_2) < \mathbb{E}_U(L_1)$. So higher *expected* epistemic value may not correspond to the agent's comparative judgments of epistemic "betterness." Because there are belief states b_1 and b_2 that can be identified with epistemic lotteries with L_1 and L_2, A1 entails that judgments about which of two belief states is more rational to adopt will therefore also not correspond to judgments about which has higher expected epistemic value. But that's precisely what is required for proper scoring rules to have normative implications.

5 Alternative Representation Theorems

Is there an alternative to von Neumann and Morgenstern that avoids all four of the axioms we target? In short, No.

Of the four axioms we have discussed, two—totality and transitivity—are necessary assumptions of *every* representation theorem that could be used to show that inaccuracy is numerical. The reason is simple: if degrees of inaccuracy are numerical, then they are always comparable because two numbers can always be compared. Comparisons of inaccuracy must be transitive for the same reason. If our arguments against these two axioms are successful, abandoning von Neumann and Morgenstern's framework is of no avail.

The independence and continuity axioms are not, however, essential to other representation theorems. In showing that length and pitch are numerical, for example, the concept of a "lottery" is irrelevant, and hence, the independence and continuity axioms above play no role. Employing one of these alternative representation theorems, however, faces difficulties. While some alternative representation theorems would guarantee uniqueness of inaccuracy up to some type of transformation, the type of transformation may not justify the meaningfulness of *expected* utilities, which EDT requires for its use of proper scoring rules.

Furthermore, alternative theorems may also require axioms that are equally implausible in the epistemic case. In the case of preference, the set of options is closed under a ternary operation $(\alpha,\cdot,\cdot)$, which forms a lottery $\langle\alpha,x,y\rangle$ for any two options x and y and any probability α. Such a structure is called a *mixture space*. By contrast, in the case of length, the set of objects is closed instead under an associative, commutative, and binary operation $+$, which represents the concatenation of two objects. Although this structure is not a mixture space, there must be some operation on belief states that behaves like concatenation if the representation theorem for length is to be employed in representing epistemic accuracy. But it is hard to see what believing that p plus believing that q could be, setting aside the difficulties any alternative representation scheme will face in satisfying the simple ordering condition entailed by totality and transitivity. Similar remarks, we maintain, apply to other representation theorems about pitch, color, and so on.[38]

Let's sum up our discussion of the continuity and independence axioms. To justify these axioms, epistemic decision theorists need to find an analog of a lottery. To steer around this requirement involves representation theorems for structures that are not mixture spaces, and therefore either include an operation that is implausible for beliefs, such as concatenation, or is otherwise too weak to carry EDT. The obvious epistemic analog of a lottery is an "intermediate" belief state. The most plausible justification for the existence of intermediate belief states is that belief itself is numerical. But the only known reasons to think belief is numerical require either (*i*) endorsing the pragmatic arguments that puritans reject or (*ii*) accepting controversial axioms of qualitative probability that render the accuracy-first defense of probabilism irrelevant.

6 The Reckoning

They criticize what you say, but they never give you credit for how loud you say it.
—Stephen Colbert (2007)[39]

Rational choice is a normative theory about *choice*. It furnishes norms to guide your *actions*, and its axioms are justified by showing that a failure to heed them leads to incoherent *behavior* judged by your own lights. The logic of sure-loss avoidance—*coherence*—concerns internal rationality, but the theory operationalizes how to assess whether acting on your commitments would undermine your interests.

EDT, by contrast, describes rational belief rather than coherent commitments. It purports to give normative standards to evaluate your judgments *vis-à-vis* the state of the world, but it provides no means for measuring credence or epistemic utility and therefore no mechanism for assessing whether your beliefs are deficient. Puritans often emphasize that severing judgments of uncertainty from preference

[38]There is a final reason that some epistemic decision theorists would have difficulty rejecting the claim that belief states form a mixture space. Namely, some also argue that a measure of inaccuracy ought to be convex or "weakly" convex, and such axioms implicitly assume that a set of belief states are closed under at least some types of mixtures. See (Joyce 1998, p. 596).

[39]EDTists follow the good Reverend Colbert's Wørd when they seek to measure inaccuracy by strength of conviction, not content.

allows you to repudiate the overly-pragmatic doctrines of traditional decision theory.[40] However, without some procedure for measuring credence or epistemic utility, EDT lacks an alternative justification for the axioms of decision-theoretic representation theorems. Calling EDT "decision theory" does nothing to establish that standard representation theorems apply to epistemic value any more than the appellation "flight" warrants thinking the mathematical model describing the motion of jumbo jets applies to hummingbirds. An argument, not equivocation, is necessary.

We recognize that it is difficult for EDT to back off from its commitment to puritanism. For to concede that quantifying epistemic loss depends on an individual's interests undercuts the central motivation for supplying probabilism with a non-pragmatic justification. To those who maintain that epistemic value cannot be relativized to a person, an impure EDT is indistinguishable from traditional decision theory and "accuracy-first" is reduced to accuracy cant. So, either EDT is a speculative metaphysical program entirely without foundations or EDT enjoys sound but pragmatic foundations and is subsumed by, rather than an alternative to, traditional decision theory.

The choice should be clear. The pragmatic, ecunemical view about the relationship between belief and preference has been tremendously fruitful. It yields an axiomatic approach to rationality that unifies research in several disciplines, including philosophy, psychology, economics, and statistics. This unified view sees pragmatic decision-making and scientific inference as lying on a continuum, in which the rationality of both the everyday decision-maker's actions and the seemingly disinterested scientist's inferences are influenced by subjective preference, but to widely differing degrees, and to widely differing standards of transparency and criticism. EDT is part of an epistemological tradition of striving for objectivity, but its puritanical creed is inconsistent with its quantitative rites.

[40]For example, Briggs (2015, pp. 630-631) claims, "Epistemic representation theorems are on more secure footing. It may be implausibly pragmatic to claim that partial beliefs are nothing over and above the preferences they (and utilities) give rise to. However, it is more plausible that there is nothing more to degrees of belief than the comparative probabilities they encode – that utilities simply measure certain features of comparative belief."

Acknowledgments (2019)

Thanks to Seamus Bradley for extensive suggestions for improving earlier drafts of this paper. We are also indebted to Colin Marshall, Sam Fletcher, and Jonathan Livengood for helping us clarify our criticism of the independence axiom. Thanks to Samir Okasha for his help in refining our arguments concerning "epistemic lotteries," Jean Baccelli for discussing this and related topics with us over the years and for extensive comments on this paper, and to both Jim Joyce and the late Isaac Levi for discussing their past and present (at the time) views on the very idea of epistemic utility. Although they do not endorse our conclusions, we would like to thank Jason Konek and Richard Pettigrew for discussions concerning alternative representation theorems, the role of representation theorems, and more features of epistemic decision theory. Finally, we are thankful of audiences at Columbia University, the University of Washington, the Munich Center for Mathematical Philosophy, the University of Bristol, and the 2015 British Society for Philosophy of Science meeting in Manchester for helpful discussions of earlier drafts of this paper.

Seattle & Frankfurt, August 27, 2019

References

Anderson, E. S. (1987). *Value in Ethics and in Economics*. Harvard University Press.

Anscombe, F. J. and R. J. Aummann (1963). A definition of subjective probability. *Annals of Mathematical Statistics 34*, 199–205.

Arlo-Costa, H. and A. Pedersen (2012). Belief and probability: A general theory of probability cores. *International Journal of Approximate Reasoning 53*(3), 293–315.

Aumann, R. (1962). Utility theory without the completeness axiom. *Econometrica 30*(3), 445–462.

Briggs, R. (2015). Corroboration and rules of acceptance. *Journal of Philosophical Logic 6*, 625–640.

Burge, T. (1993). Content preservation. *Philosophical Review 102*(4), 457–488.

Chang, R. (2002). The possibility of parity. *Ethics 112*, 659–688.

Colbert, S. (2007). The O'Reilly Factor. January 18, 2007 Episode. Fox News, Inc.

D'Agostino, M. and C. Sinigaglia (2010). Epistemic accuracy and subjective probability. In M. Suárez, M. Dorato, and M. Rédei (Eds.), *EPSA Epistemology and Methodology of Science*, pp. 95–105. Springer.

de Finetti, B. (1931). On the subjective meaning of probability. *Fundamenta Mathematica 17*, 298–329. Republished in Monari, P. and Cocchi, D. (Eds.), *Induction and Probability*, Bologna: CLUEB, 1992.

de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré 7*(1), 1–68.

de Finetti, B. (1974). *Theory of Probability* (1990 ed.), Volume I. New York: John Wiley.

Dubra, J., F. Maccheroni, and E. A. Ok (2004). Expected utility theory without the completeness axiom. *Journal of Economic Theory 115*(1), 118–133.

Easwaran, K. (2015). Decision theory without representation theorems. *Philosophers Imprint*.

Easwaran, K. and B. Fitelson (2015). Accuracy, coherence, and evidence. In T. S. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology*, Volume 5, pp. 61–96. Oxford: Oxford University Press.

Fishburn, P. C. (1979). *Utility Theory*. Kreiger Publishing Company.

Fishburn, P. C. (1986). The axioms of subjective probability. *Statistical Science 1*(3), 335–358.

Fishburn, P. C. (1991). Nontransitive preferences in decision theory. *Journal of Risk and Uncertainty 4*(2), 113–134.

Gibbard, A. (2007). Rational credences and the value of truth. In T. S. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology*, Volume 2, pp. 143–164. Oxford University Press.

Gilboa, I. (2009). *Theory of Decision Under Uncertainty*. Cambridge University Press.

Greaves, H. and D. Wallace (2006). Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind 115*, 607–632.

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

Joyce, J. M. (2005). How probabilities reflect evidence. *Philosophical Perspectives 19*, 153–178.

Joyce, J. M. (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In *Degrees of belief*, pp. 263–297. Springer.

Kelley, M. (2019). *Accuracy Dominance on Infinite Opinion Sets*. Master of Arts thesis, University of California, Berkeley.

Konek, J. and B. Levinstein (2016). The foundations of epistemic decision theory. *Mind*. Unpublished Manuscript.

Kraft, C. H., J. W. Pratt, and A. Seidenberg (1959). Intuitive probability on finite sets. *The Annals of Mathematical Statistics 30*(2), 408–419.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971). *Foundations of Measurement Volume I: Additive and Polynomial Representations* (Dover Publications, 2006 ed.). Academic Press.

Leitgeb, H. and R. Pettigrew (2010). An objective justification of Bayesianism I: Measuring accuracy. *Philosophy of Science 77*(2), 201–235.

Levi, I. (1963). Corroboration and rules of acceptance. *The British Journal for the Philosophy of Science 13*, 307–313.

Levi, I. (1967). *Gambling with Truth*. Cambridge, MA: MIT Press.

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica 24*(2), 178–191.

Luce, R. D. (1967). Sufficient conditions for the existence of a finitely additive probability measure. *The Annals of Mathematical Statistics 38*, 780–786.

Mayo-Wilson, C. and G. Wheeler (2016). Scoring imprecise credences: A mildly immodest proposal. *Philosophy and Phenomenological Research 93*(1), 55–78.

Oddie, G. (1997). Conditionalization, cogency, and cognitive value. *The British Journal for the Philosophy of Science 48*, 533–541.

Okasha, S. (2011). Theory choice and social choice: Kuhn versus arrow. *Mind 120*(477), 83–115.

Pedersen, A. P. (2014). Comparative expectations. *Studia Logica 102*(4), 8111–848.

Pettigrew, R. (2012). Accuracy, chance, and the Principal Principle. *Philosophical Review 121*(2), 241–275.

Pettigrew, R. (2013a). Accuracy and evidence. *Dialectica 67*(4), 579–596.

Pettigrew, R. (2013b). Epistemic utility and norms for credences. *Philosophy Compass 8*(10), 897–908.

Pettigrew, R. (2015). Risk, rationality and expected utility theory. *Canadian Journal of Philosophy 45*(5-6), 798–826.

Pettigrew, R. (2016a). *Accuracy and The Laws of Credence*. Oxford: Oxford University Press.

Pettigrew, R. (2016b). Accuracy, risk, and the principle of indifference. *Noûs 92*(1), 35–59.

Predd, J. R., E. Seringer, H. Lieb, D. Osherson, V. Poor, and S. Kulkami (2009). Probabilistic coherence and proper scoring rules. *IEEE Transactions on Information Theory 55*(10), 4786–4792.

Ramsey, F. P. (1926). Truth and probability. In R. B. Braithwaite (Ed.), *The Foundations of Mathematics and Other Logical Essays*, pp. 156–198, 1931. London: Kegan, Paul, Trench & Company.

Ramsey, F. P. (1990 (1929)). Probability and partial belief. In D. H. Mellor (Ed.), *Philosophical Papers*, pp. 95–96. Cambridge University Press.

Rosenkrantz, R. (1981). *Foundations and Applications of Inductive Probability*. Atascadero: Ridgeview Publishing Company.

Savage, L. J. (1954). *Foundations of Statistics*. New York: Wiley.

Schervish, M. J., T. Seidenfeld, and J. B. Kadane (2014). Dominating countably many forecasts. *The Annals of Statistics 42*(2), 728–756.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (1995). A representation of partially ordered preferences. *The Annals of Statistics 23*(6), 2168–2217.

Sen, A. K. (1971). Choice functions and revealed preference. *The Review of Economic Studies 38*(3), 307–317.

Sen, A. K. (1973). Behaviour and the concept of preference. *Economica 40*(159), 241–259.

Sen, A. K. (1982). *Choice, Welfare and Measurement*. Cambridge, MA: MIT Press.

Smith, C. A. B. (1961). Consistency in statistical inference (with discussion). *Journal of the Royal Statistical Society 23*, 1–37.

Sousa, R. B. D. (1974). The good and the true. *Mind 83*(332), 534–551.

Vicig, P. (2016). A note on the equivalence of coherence and constrained coherence. *Minds and Machines 26*(3).

von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.

Wheeler, G. (2018). Bounded rationality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018 ed.). Metaphysics Research Lab, Stanford University.