The Qualitative Grounds of Quantitative Probability

Edward Elliott

Department of Philosophy, University of Notre Dame, 100 Malloy Hall, Notre Dame, 46556, IN, United States.

Contributing authors: eelliot2@nd.edu;

Abstract

What is the qualitative basis for deriving meaningful psychological information from probabilistic representations of confidence within decision theory? Conditions on two binary relations—a preference relation $\succsim$ and a comparative confidence relation $\succsim_{\boldsymbol{c}}$—are shown to imply the existence of a probability function and a utility function that jointly represent $\succsim$ under the expected utility rule, where the probability function is unique and also represents $\succsim_{\boldsymbol{c}}$. It is also shown that, under certain conditions, these probabilities contain more *meaningful* information than can be extracted from the confidence relation on its own. The upshot is that there's more to what's meaningful in decision-theoretic representations of confidence than can generally be explained by reference to confidence relations alone.

Keywords: Qualitative probability, Comparative confidence, Representation theorem, Ratio information, Meaningfulness

1 Introduction

This paper presupposes a realist interpretation of (subjective) expected utility theory. In particular, I assume that the probability functions we find within expected utility theory are best thought of as numerical representations of *confidence*, where this is understood to be neither conceptually nor ontologically reducible to preferences. Instead, we are to think of confidence as a genuine psychological quantity, one posited to play a central role in both causally explaining and normatively constraining the shape of an agent's preference ordering.

Taking this kind of realism for granted, some questions arise in relation to how the probabilistic representation of confidence ought to be understood. After all, the mere assertion that probabilities represent confidence does not settle *what* is meaningful in such representations, nor *why*. Start with the *what*-question, and suppose that the function p assigns probabilities 0.99 and 0.33 to events E and F respectively. Obviously, p represents more confidence in E than F. But is there any more to it? For instance, the difference between 0.99 and 0.33 is twice the difference between 0.33 and 0—is that relation supposed to track something real? Similarly, 0.99 is three times 0.33—is *that* relation supposed to track something real? Are these facts about ratios meaningful, or are they mere artifacts of our representational conventions? To put it in more familiar terms, the *what*-question concerns the kind of scale on which confidence is being represented: is it an ordinal scale, an interval scale, or a ratio scale?

Those are far from the only scale types, of course, but for present purposes we do not need a full taxonomy. (For that, see Narens [1].) My interest lies specifically in the ratio scale hypothesis. In fact, it is quite straightforward to show that probability ratios *are* meaningful within expected utility theory, as it's typically formulated, at least under certain conditions. But then there's the question of *why*. Certainly, the fact that $0.99 = 3 \times 0.33$ doesn't justify the claim that the *three times* ratio is meaningful with respect to probabilistic representations of confidence, any more than it would justify the claim that 99° F ($\approx 37°$ C) is three times hotter than 33° F ($\approx 0.5°$ C). Rather, if probability ratios are meaningful, then the explanation for their meaningfulness lies not in the numbers themselves but in some qualitative feature of the psychological system those numbers are supposed to represent. So what is it about *confidence*, as it's portrayed within expected utility theory, that explains why probability ratios are meaningful within that theory when they are?

This paper presents a representation theorem with the aim of shedding some light on these questions. Assuming a realist interpretation of expected utility theory throughout, I will use this theorem to argue for the following two-part thesis:

Preference-Grounded Ratios. There are conditions under which (i) two measures of confidence are equivalent in meaning just in case they agree on all ratio information, and moreover (ii) the explanation for this fact relates at least in part to the relationship between confidence and preference.

The first part is unlikely to be very controversial.[1] The second part, however, stands in contrast to the historically popular idea that probability ratios are meaningful by virtue of the structure of the *comparative confidence* relations they represent—that is, purely ordinal comparisons of relative confidence such as "has more confidence in E than in F" and "has the same amount of confidence in E and F". I will discuss how these comparativist explanations work in due course; what's important for now is that they require no direct reference to preferences, and therefore cast doubt on the second part of Preference-Grounded Ratios.[2]

[1] Though cf. Joyce [2]: 416–9. Joyce argues, given a slight reformulation of the expected utility rule, that probabilities ought to be seen as representing confidence on an interval scale—so ratios of differences are meaningful, but not ratios *simpliciter*.

[2] The idea that we can explain why subjective probability ratios are meaningful in terms of comparative confidence relations has an extensive history—see de Finetti [3], Koopman [4], [5], Savage [6], Villegas [7], Luce [8], Krantz et al. [9]: ch. 5, Fine [10]: ch. 2, Suppes and Zanotti [11], [12], Suppes [13], and Zynda [14];

There are six sections to follow. Section 2 contains the relevant formal preliminaries, while Section 3 discusses the conceptual background in more detail—focusing especially on the key things required of a representation theorem if it is to be applied in support of Preference-Grounded Ratios in the way that I have in mind. Section 4 then presents the theorem in full. Section 5 explicates the notion of *meaningfulness*, and on that basis argues for part (i) of my thesis; and finally, Section 6 and Section 7 together comprise the argument for part (ii).

But before going any further, two clarifications are warranted. First, my topic concerns the meaningfulness of probability ratios specifically within the context of expected utility theory under a realist interpretation. What I have to say won't much depend on whether that theory is better understood *descriptively* (e.g., as describing the typical relationship between confidence and preference for real people) or *normatively* (e.g., as characterizing what that relationship ideally ought to be). I take no stance on such matters here, though for ease of expression I have adopted language befitting a descriptive interpretation throughout. Second, my thesis is not that probability ratios are *always* meaningful, nor necessarily that the explanation for their meaning *always* requires reference to preferences. It is consistent with Preference-Grounded Ratios that different explanatory strategies may be appropriate under different circumstances, or for different theoretical purposes. The contention is simply that there *are* circumstances in which meaningful probability ratios are best explained at least in part via the relationship between confidence and preference under conditions of uncertainty.

2 Formal preliminaries

We adopt an analytic framework in the style of Savage [6]. To start, there are two primitive sets ($\mathcal{S}$ and $\mathcal{C}$) and two derived sets ($2^{\mathcal{S}}$ and $\mathcal{A}$):

- $\mathcal{S}$ is a non-empty set of *states*, with typical element s.
- $\mathcal{C}$ is a non-empty set of *consequences*, with typical elements $\alpha, \beta, \gamma, \delta$.
- $2^{\mathcal{S}}$ is the set of *events* (i.e., sets of states), with typical elements E, F.
- $\mathcal{A} \subseteq \mathcal{C}^{\mathcal{S}}$ is the set of *simple acts* (i.e., finitely-valued functions from states to consequences), with typical elements f, g, h.

In Section 4, I will assume that both $\mathcal{S}$ and $\mathcal{C}$ are finite (and so $\mathcal{A}$ just is $\mathcal{C}^{\mathcal{S}}$), and I'll discuss this assumption further in Section 6. But for now we need make no assumptions regarding the size of either set.

We adopt the following notational conventions. First, for any event E, the complement of E with respect to $\mathcal{S}$ is denoted $\overline{E}$; and s will be used ambiguously for both the state and the singleton event that contains it. Next, we use fEg for the mixture of acts f and g on event E. In other words, fEg denotes the act h such that $h(s) = f(s)$ for all states s in E and $h(s) = g(s)$ otherwise. Furthermore, we let $f_E = \alpha$ whenever

for more recent discussions, see Suppes [15], DiBella [16], Stefánsson [17], Konek [18], Elliott [19] and Neth [20]. Such relations have been given a variety of overlapping names across several literatures, including but not limited to *comparative probabilities*, *qualitative probabilities*, *intuitive probabilities*, *qualitative confidence*, *comparative beliefs*, and *belief orderings*.

$f(s) = \alpha$ for all s in E; and we identify each consequence with its corresponding constant act, so $f = \alpha$ whenever $f_{\mathcal{S}} = \alpha$. Given the above, note in particular that $\alpha E \beta$ designates the *binary act* f such that $f_E = \alpha$ and $f_{\overline{E}} = \beta$.

In addition, there are two relations that will be central to the discussion:

- $\succsim$ is a binary *preference relation* on $\mathcal{A}$.
- $\succsim_c$ is a binary *confidence relation* on $2^{\mathcal{S}}$.

We take the indifference $\sim$ and strict preference $\succ$ relations to be the symmetric and asymmetric parts of $\succsim$ respectively. Furthermore, we say that $\succsim$ is *non-trivial* whenever the asymmetric part is non-empty; and it is a *weak order* whenever it is complete (i.e., $f \succsim g$ or $g \succsim f$ for all f, g) and transitive. The same goes for $\succsim_c$, $\sim_c$ and $\succ_c$, *mutatis mutandis*.[3]

Next, we refer to any real-valued function on the set of consequences as a *utility function*. Similarly, we refer to any real-valued function on the set of events as a *confidence function*. The flexibility in what counts as a confidence function is intentional, as we want minimal presuppositions regarding how confidence is being numerically represented. By contrast, a *probability function* is to be defined strictly as any confidence function that satisfies the usual properties:

- $c(\mathcal{S}) = 1$ (*normalization*)
- $c(E) \geq 0$ (*non-negativity*)
- $E \cap F = \varnothing$ implies $c(E \cup F) = c(E) + c(F)$ (*additivity*)

For all that follows, p will be reserved for probability functions, whereas c will be used for confidence functions that may satisfy some, all, or none of these three properties.

Finally, note that every simple act induces at least one (and typically more than one) finite partition $\mathcal{E}_f$ of $\mathcal{S}$ such that f is constant with respect to each E in $\mathcal{E}_f$. Given that, let the valuation $V_{c,u}$ be a real-valued function on $\mathcal{A}$ defined relative to some (not necessarily probabilistic) confidence function c and utility function u such that for all f and any corresponding choice of consequence partition $\mathcal{E}_f$,

$$V_{c,u}(f) = \sum_{E \in \mathcal{E}_f} c(E) \cdot u(f_E).$$

Given this, say that $V_{c,u}$ *represents* $\succsim$ just in case

$$f \succsim g \Leftrightarrow V_{c,u}(f) \geq V_{c,u}(g).$$

Where c is normalized, non-negative and additive, this is obviously just ordinary expected utility maximization. I take it as given that the central principle of expected utility theory, under any interpretation realist or otherwise, is that $\succsim$ maximizes expected utility relative to at least one probability function.

[3] Keep in mind that $\succsim_c$ has here been defined specifically as a binary relation on $2^{\mathcal{S}}$; in other words, it is an *unconditional* confidence relation. In Section 6, we will also consider *conditional* confidence relations that hold between ordered pairs of events, but until then we restrict attention to the simpler unconditional relations.

3 Conceptual motivations

The intended argument for Preference-Grounded Ratios requires a theorem—let's call it the PGR theorem—with certain properties that set it somewhat apart from other representation theorems such as Savage's [6]. As such, I want to first explain what those properties are and why they're important. To that end, it will be useful to set the PGR theorem in contrast to Savage's classic result.

Without going into specifics, Savage showed that if an agent's preferences satisfy certain conditions—he called them *postulates*—then they will maximize expected utility relative to exactly one probability function p. It has often been remarked, however, that nothing about Savage's theorem by itself implies that this probability function is therefore a representation of the agent's confidence.[4] You might think, for instance, that if $\succsim_c$ is the agent's comparative confidence ordering, then p should at minimum *represent* $\succsim_c$ in the sense that

$$p(E) \geq p(F) \;\Leftrightarrow\; E \succsim_c F.$$

But for all that Savage's theorem directly tells us, p might have no particular connection to confidence whatsoever. Indeed, since the theorem only directly concerns the representation of preferences, p need not correspond to anything at all beyond certain patterns in an appropriately structured preference relation.

Savage's position on these matters was roughly that confidence—he called it *personal probability*—ought to be characterized in terms of preferences.[5] In particular, it's easy to see that if $\alpha \succ \beta$, then $\succsim$ maximizes expected utility relative to p only if

$$p(E) \geq p(F) \;\Leftrightarrow\; \alpha E \beta \succsim \alpha F \beta.$$

In light of this, Savage proposed to define the agent's comparative confidence relation in terms of their preferences like so:

$$E \succsim_c F \;\Leftrightarrow\; \forall \alpha, \beta \big(\alpha \succ \beta \;\Rightarrow\; \alpha E \beta \succsim \alpha F \beta\big).$$

With that in place, it follows immediately that p represents $\succsim_c$ whenever $\succsim$ maximizes expected utility relative to p.

Now there are two ways one might think of Savage's 'definition'. In its crudest form, we can treat it as a proposed ontological reduction—thus, p represents $\succsim_c$ *because* it represents certain patterns in $\succsim$, and that's all that $\succsim_c$ really is. This would be plainly inconsistent with the realist perspective I assumed at the outset. However, there is another way to see things. Suppose we take $\succsim_c$ to be an independent primitive in our framework, alongside the usual preference relation $\succsim$. That is, we do not try to define $\succsim_c$ in terms of $\succsim$, but instead posit $\succsim_c$ as its own thing with an interpretation that need not be tied in any particular way to $\succsim$. Given that, we can then re-construe Savage's

[4] See, for instance, Christensen [21], 358–364; Eriksson and Hájek [22], 196–202; Hájek [23], 803–4; Meacham and Weisberg [24], 644–5; Titelbaum [25], 298–9; Zynda [14].

[5] See Savage [6], especially ch. 3; and [26], 594–597. It should be noted that Savage's definition of 'personal probability' was intended in the first instance for an 'idealized *homo economicus*—and therefore, with some approximation, a real person'; see Savage [27], 783.

'definition' as a substantive constraint on how comparative confidence *correlates* with certain patterns in the preference ordering, to be added as a further condition of the theorem. Indeed, that this correlation holds ends up being a necessary condition for the truth of expected utility theory under a realist interpretation.

The point here is that we require a *dual-primitive* theorem, one that treats preference and comparative confidence on equal footing. Other such theorems exist, with the best known being Joyce's [28].[6] However, there is a difference that's worth highlighting. In any dual-primitive theorem, we can distinguish three kinds of conditions: (i) those that directly concern only preference, (ii) those that directly concern only comparative confidence, and (iii) those that specify how the two relations are connected. Joyce makes liberal use of the second type, with the result being that his comparative confidence relations are guaranteed to have a unique probabilistic representation *regardless* of how they relate to preferences.[7] By contrast, the PGR theorem employs minimal conditions of type (ii)—not because doing so is especially desirable from a realist perspective, but because this is what best serves my argumentative purposes. Specifically, I want to argue that sometimes what's meaningful in a probability function cannot be understood without reference to preferences. To that end, it is better to minimize any direct constraints on $\succsim_c$—lest one be accused of smuggling ratios in through structure posited directly on the confidence relation.

This point is directly related to the next important contrast with Savage's theorem that I want to highlight. In order to show that preferences are (sometimes) required to explain meaningful probability ratios in expected utility theory, one must isolate conditions consistent with that theory under which some preference-based explanation *can* apply while other explanations *cannot*. This rules out Savage's theorem, among others, since the method by which Savage delivers unique probabilities makes central use of conditions that preclude precisely this kind of explanatory differentiation. In particular, Savage's conditions hold only when there are explanations of confidence ratios available that depend only on the structure of the confidence relation, and therefore require no direct reference to preferences.

This bears some explanation. To start, note that while every probability function represents exactly one confidence relation, some confidence relations will be represented by more than one probability function. As such, say that p is *ordinally determined* just in case, if it represents $\succsim_c$, then it is the only probability function that does so; otherwise, p is ordinally *under*determined. Given that, the strategy of Savage's proof can be broken into three steps:

1. Define $\succsim_c$ in terms of $\succsim$.
2. Show that p is ordinally determined by $\succsim_c$.
3. Derive a utility function u (unique up to a positive affine transformation) such that $V_{p,u}$ represents $\succsim$.

The important part for our purposes is step 2, since it means that a preference relation will satisfy all of Savage's postulates only if it maximizes expected utility relative to an ordinally determined probability function.

[6] See also Jeffrey [29], 77–78, for an earlier statement of the idea.
[7] See Joyce [28], 137–139, 233–240.

Why does that rule out Savage's theorem for the purposes of my argument? Because p is ordinally determined whenever the confidence relation it represents has *exactly* the kind of structure that suffices to explain all of the ratio information contained within p. This follows from a well-known result by Suppes and Zanotti [11]. Let $E^{\mathrm{i}} : \mathcal{S} \to \{0, 1\}$ be the indicator function for event E relative to $\mathcal{S}$; i.e.

$$E^{\mathrm{i}}(s) = \begin{cases} 1 & \text{if } s \in E, \\ 0 & \text{if } s \notin E. \end{cases}$$

Next, an *extended indicator function* is a non-negative integer combination of indicator functions. That is, for any positive integer n, events $E_1, \ldots, E_n$, and non-negative integers $k_1, \ldots, k_n$, $X : \mathcal{S} \to \mathbb{N}_0$ is an extended indicator function just in case

$$X(s) = \sum_{j=1}^{n} k_j \cdot E_j^{\mathrm{i}}(s).$$

Finally, let $\mathcal{A}^\star$ be the *algebra of extended indicator functions generated by* $\mathcal{S}$—defined as the smallest semigroup under functional pointwise addition that contains the indicator functions for all subsets of $\mathcal{S}$.[8] Then we have:

Theorem 1 (Suppes and Zanotti [11]; Suppes [15]) *Let $\succsim^\star$ be a binary relation over the algebra $\mathcal{A}^\star$ of extended indicator functions generated by $\mathcal{S}$. Then $\succsim^\star$ satisfies, for all $X, Y, Z \in \mathcal{A}^\star$,*

SZ1. $\succsim^\star$ *is a weak order,*

SZ2. $X \succsim^\star \varnothing^{\mathrm{i}}$,

SZ3. $\mathcal{S}^{\mathrm{i}} \succ^\star \varnothing^{\mathrm{i}}$,

SZ4. $X \succsim^\star Y \Leftrightarrow X + Z \succsim^\star Y + Z$,

SZ5. $X \succ^\star Y$ *only if* $nX \succ^\star k\mathcal{S}^{\mathrm{i}} \succ^\star nY$ *for some positive integers* k, n,

if and only if there is a real-valued expectation function $\mathbb{E}$ on $\mathcal{A}^\star$ such that for all $X, Y \in \mathcal{A}^\star$,

E1. $\mathbb{E}(X) \geq 0$,

E2. $\mathbb{E}(\varnothing^{\mathrm{i}}) = 0$,

E3. $\mathbb{E}(\mathcal{S}^{\mathrm{i}}) > 0$,

E4. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$,

E5. $X \succsim^\star Y \Leftrightarrow \mathbb{E}(X) \geq \mathbb{E}(Y)$.

Furthermore, if $\mathbb{E}$ and $\mathbb{E}'$ are any two expectation functions on $\mathcal{A}^\star$ satisfying E1–E5*, then there is a positive real number* a *such that $\mathbb{E} = \mathrm{a}\mathbb{E}'$; that is, $\mathbb{E}$ is unique up to a positive scalar transformation.*

---

[8] That is, for all events E and extended indicator functions X, Y: (i) $E^{\mathrm{i}} \in \mathcal{A}^\star$; (ii) if $X, Y \in \mathcal{A}^\star$ then $X + Y \in \mathcal{A}^\star$; and (iii) nothing else is in $\mathcal{A}^\star$. Here, $+$ is defined pointwise in the usual way: $(X + Y)(s) \coloneqq X(s) + Y(s)$, for all s.

As Suppes and Zanotti point out, the properties SZ1–SZ5 are of the same form as those used to characterize positive extensive measurement structures (cf. Krantz et al. [9]). What their result tells us, then, is that a comparative confidence ordering has a unique probabilistic representation just in case it has—or determines under an appropriate extension—the very same kind of structure used to underwrite the ratio-scale measurement of extensive quantities such as length, mass, and duration. More precisely, suppose we first induce $\succsim^\star$ on the space of (ordinary) indicator functions using $\succsim_c$:

$$E \succsim_c F \Leftrightarrow E^{\mathrm{i}} \succsim^\star F^{\mathrm{i}}.$$

It is then a corollary of Suppes and Zanotti's theorem that a complete $\succsim_c$ ordering has a unique probabilistic representation if and only if there is a unique extension of $\succsim^\star$ to the full algebra of extended indicator functions generated by $\mathcal{S}$ that satisfies SZ1–SZ5.[9]

To get a picture of how this works, imagine that $\succsim_c$ represents the relative *size* of events—so $E \sim_c F$ means that E is the same size as F, while $E \succ_c F$ means that E is bigger than F. We assume that any event is at least as big as $\varnothing$, which has zero size; and that sizes combine additively in the usual way. Let $\mathcal{S} = \{s_1, s_2, s_3, s_4\}$, and label non-singleton events via the indices of the states they contain in the obvious way. Now imagine two ways that this ordering might go:

$$(a)\ \mathcal{S} \succ_c \begin{matrix} E_{123} \\ E_{124} \\ E_{134} \end{matrix} \succ_c \begin{matrix} E_{12} \\ E_{13} \\ E_{14} \end{matrix} \succ_c \begin{matrix} s_1 \\ E_{234} \end{matrix} \succ_c \begin{matrix} E_{23} \\ E_{24} \\ E_{34} \end{matrix} \succ_c \begin{matrix} s_2 \\ s_3 \\ s_4 \end{matrix} \succ_c \varnothing$$

and

$$(b)\ \mathcal{S} \succ_c \begin{matrix} E_{123} \\ E_{124} \\ E_{134} \end{matrix} \succ_c \begin{matrix} E_{12} \\ E_{13} \\ E_{14} \\ E_{234} \end{matrix} \succ_c \begin{matrix} s_1 \\ E_{23} \\ E_{24} \\ E_{34} \end{matrix} \succ_c \begin{matrix} s_2 \\ s_3 \\ s_4 \end{matrix} \succ_c \varnothing$$

In each case, the ordering agrees with exactly one probability function, as represented by Figure 1 (a) and (b). In the first case, observe in particular that the state space $\mathcal{S}$ can be partitioned into two equal-sized parts (s_1 and E_{234}), which implies that each part is $\frac{1}{2}$ the size of $\mathcal{S}$. Likewise, E_{234} can in turn be partitioned into three equal-sized parts (s_2, s_3 and s_4), which are therefore all $\frac{1}{6}$ the size of $\mathcal{S}$. In this manner, the relative sizes of each state—and therefore the relative sizes of all events composed from those states—can be determined simply by iteratively breaking every event down into equal-sized parts.

For (b), by contrast, there is no way to partition $\mathcal{S}$ into equal-sized parts. However, the same kind of 'equal-parts' reasoning implies that s_1 is the same size as E_{23}, which is twice the size of s_2, s_3 and s_4. In light of that, one way to understand the Suppes and Zanotti construction is to imagine what would happen if we could 'duplicate' any event

[9] See Suppes and Zanotti [11], 435, though note that they slightly misstate this result. They claim that a complete $\succsim_c$ has a unique probabilistic representation just in case there *is* an extension to some $\succsim^\star$ that satisfies SZ1–SZ5. A simple counterexample is to let $\mathcal{S} = \{s_1, s_2\}$ and then consider the ordering $\mathcal{S} \succ_c s_1 \succ_c s_2 \succ_c \varnothing$. There are many ways to extend $\succsim_c$ to $\succsim^\star$ so as to satisfy SZ1–SZ5. By Theorem 1, these different extensions determine different $\mathbb{E}$ (unique up to a positive scalar transformation), which in turn induce different probability functions that all represent $\succsim_c$. The fix is immediate: $\succsim_c$ has a unique probabilistic representation whenever it has a *unique* extension.

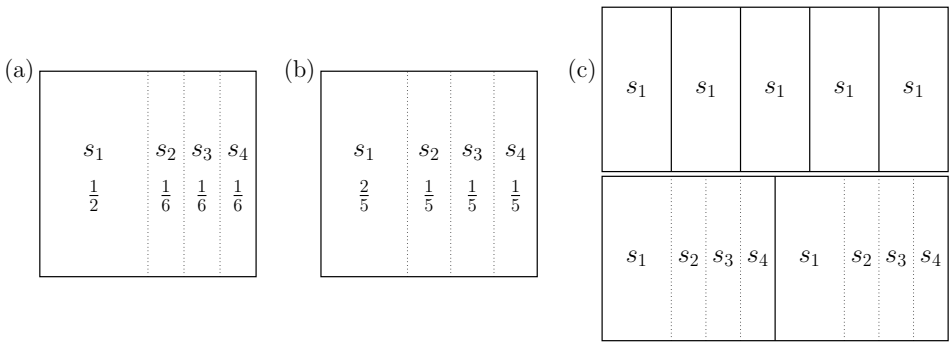

Fig. 1 Relative-size representations of two comparative confidence orderings. Panels (a) and (b) represent the normalized state sizes corresponding to the two orderings; panel (c) illustrates the extended-indicator construction used to determine the sizes in (b).

as many times as we like and combine them together in different ways. In particular, think of the extended indicator function $5 \cdot s_1^{\mathrm{i}}$ as representing the combination of five 'duplicates' of s_1, and likewise $2 \cdot \mathcal{S}^{\mathrm{i}}$ represents the combination of two 'duplicates' of $\mathcal{S}$. Now given what $\succsim_c$ tells us about the relative sizes of each state, it follows that $5 \cdot s_1^{\mathrm{i}}$ would have to be the same size as $2 \cdot \mathcal{S}^{\mathrm{i}}$, as pictured in Figure 1 (c). As such, s_1 must be $\frac{1}{5}$ the size of $2 \cdot \mathcal{S}^{\mathrm{i}}$—or equivalently, $\frac{2}{5}$ the size of $\mathcal{S}$—and consequently, s_2, s_3 and s_4 must each be $\frac{1}{5}$ the size of $\mathcal{S}$.

To summarize, the Suppes and Zanotti result tells us that whenever p is ordinally determined, the total pattern of ratios encoded by p is meaningful; and in that case, there is a satisfactory measurement-theoretic explanation of those ratios which depends solely on the structure of $\succsim_c$. (Stronger: it tells us that this particular explanation is available *only if* p is ordinally determined.) This rules out, for the purposes of my argument, any representation theorem that delivers only ordinally determined probabilities. That includes Savage's theorem, ultimately due to his chosen order of construction:

$$\textit{preferences} \rightarrow \textit{unique probabilities} \rightarrow \textit{interval-scale utilities}$$

For this reason, the PGR theorem is based on an order of construction similar to the one originally employed by Ramsey [30]; namely:

$$\textit{preferences} \rightarrow \textit{interval-scale utilities} \rightarrow \textit{unique probabilities}$$

In more detail: first we construct an 'equal-spaced' utility function over consequences; then we use those utilities to define 'exchange rates' between events that suffice to pin down a unique probability function. The benefit of this method is that it bypasses any need to posit an 'equal-parts' structure in $\succsim_c$, essentially because most of the heavy lifting is off-loaded onto the 'equal-spaced' structure in $\succsim$ over $\mathcal{C}$. This permits a certain flexibility in the kinds of probabilities that can thereby be derived—and in particular, it leaves plenty of room for ordinal underdetermination.

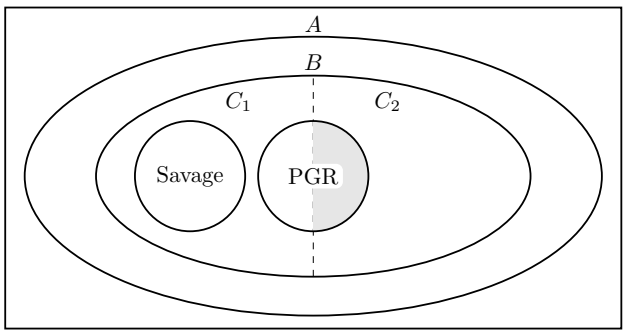

Fig. 2 Classes of preference orderings with expected utility representations. A contains those that maximize expected utility relative to at least one probability function; B, those for which that probability function is unique; and C_1 and C_2 partition B according to whether the unique probability function is ordinally determined. The shaded region is $\text{PGR} \cap C_2$.

The final contrast I want to highlight is perhaps the most important: the PGR theorem is not intended to provide an *axiomatization* of expected utility theory. This is especially relevant to how the conditions on $\succsim$ and $\succsim_c$ ought to be evaluated. If the goal were to show how expected utility theory might be reduced to a small set of empirically plausible or rationally compelling axioms, then it would make sense to evaluate these conditions according to how plausible or compelling they happen to be.[10] In the theory of measurement, however, representation theorems are traditionally employed to help explain *what* is meaningful in the representation of a quantity, and *why*. In a similar vein, my goal is to show how meaningful probability ratios arise within expected utility theory, *specifically under circumstances that enable us to differentiate between explanatory hypotheses that would otherwise be indistinguishable.*

As such, the reader is forewarned against interpreting the "axioms" of the PGR theorem either as empirical generalizations or as proposed canons of rationality—they need be neither, and some of them clearly aren't. The crucial point for my argument is that the conditions characterize possible ways an expected utility maximizer *could* in principle be. Figure 2 provides an illustration of the point. Suppose that A contains all preference orders that maximize expected utility relative to at least one p, and B is the proper subset of A in which that p is unique. B then splits into C_1, where p is ordinally determined; and C_2, where it isn't. Any theorem, such as Savage's, that is confined wholly to C_1 is thereby irrelevant for my argument. By contrast, the PGR theorem characterizes a subset of B that intersects C_1 and C_2—and for my argument, what matters most is just that the shaded region ($\text{PGR} \cap C_2$) is non-empty.

Why does that matter? Because for any preferences in the shaded region, we can use the PGR theorem to show that $\succsim$ maximizes expected utility relative to a unique probability p, even though p is ordinally underdetermined. Since the PGR theorem *also*

[10] On this basis, for instance, Savage's 'event richness' postulate P6 has been criticized for being neither descriptively plausible nor rationally compelling—e.g., Suppes [31], 162–163. P6 asserts that if $f \succ g$, then for all α, β there is a partition $\{E_1, \ldots, E_n\}$ of $\mathcal{S}$ such that $\alpha E_i f \succ g$ and $f \succ \beta E_i g$. P6 is the main reason that Savage's probability functions are ordinally determined, and the reason that Savage's postulates are satisfied only if the state space is infinite. It is noteworthy, therefore, that P6 is usually thought of as a useful simplifying assumption rather than a plausible requirement of rationality—see, e.g., Shafer [32], 468; Gilboa [33], 103, 123; Abdellaoui and Wakker [34], 2.

implies in those cases that p represents $\succsim_c$ and that the ratio information it encodes is meaningful (Section 5), it follows that what's meaningful about those ratios cannot be fully explained in terms of $\succsim_c$ (Section 6). Some other explanation is required, and after canvassing some other options, I'll argue that the best explanation in the relevant circumstances is one that involves preferences (Section 7).

4 Representation theorem

Assume that $\mathcal{S}$ and $\mathcal{C}$ are finite. There are eight conditions labeled C1–C8, the first two of which are:

C1. $\succsim$ is a weak order
C2. $\alpha \succ \beta$ implies $\alpha E \beta \succsim \alpha F \beta \;\Leftrightarrow\; E \succsim_c F$

C1 is necessary for any $V_{c,u}$ to represent $\succsim$, regardless of what c and u happen to be. C2 is just Savage's definition of $\succsim_c$ as discussed earlier, and serves as the main route by which confidence is tied to preference.

The next three conditions are what give rise to an equal-spaced utility function, in which adjacent points on a finite utility scale are spaced equally far apart. The basic idea originates from Davidson and Suppes [35], whose representation theorem also establishes an equal-spaced utility function—though here the strategy has been simplified and reconstructed for the more general Savage-style setting. First we can define the integer-valued function Δ on $\mathcal{C} \times \mathcal{C}$ recursively like so:

$$\Delta(\alpha, \beta) = \begin{cases} 0 & \text{if } \alpha \sim \beta, \\ 1 & \text{if } \alpha \succ \beta \text{ and there's no } \gamma \in \mathcal{C} \text{ such that } \alpha \succ \gamma \succ \beta, \\ \mathrm{n} & \text{if there is a } \gamma \text{ such that } \Delta(\alpha, \gamma) = \mathrm{n} - 1 \text{ and } \Delta(\gamma, \beta) = 1, \\ -\mathrm{n} & \text{if } \Delta(\beta, \alpha) = \mathrm{n}. \end{cases}$$

In other words, $\Delta(\alpha, \beta)$ denotes the number of successive 'steps' (from better to worse) to get from α to β in the preference ordering over consequences. Given that,

C3. There is an event E such that $E \sim_c \overline{E}$
C4. If $E \sim_c \overline{E}$, then $\alpha E \beta \sim \gamma E \delta$ implies $\Delta(\alpha, \gamma) = \Delta(\delta, \beta)$
C5. $\Delta(\alpha, \alpha') = \Delta(\beta, \beta')$ implies $\alpha E f \succsim \beta E g \;\Leftrightarrow\; \alpha' E f \succsim \beta' E g$

Henceforth, I will use H (or *half-probability event*) for an event that satisfies C3. In those terms, C3 states that at least one half-probability event exists. The sole purpose of this condition is to ensure that C4 is non-trivially satisfied—in which case C4 reflects the fact that if adjacent consequences are spaced equally in utility, then the difference between α and γ is the same as the difference between δ and β only if there's the same number of 'steps' between them:

$$u(\alpha) - u(\gamma) = u(\delta) - u(\beta) \;\Rightarrow\; \Delta(\alpha, \gamma) = \Delta(\delta, \beta).$$

To see this, observe that p represents $\succsim_c$ only if $p(H) = p(\overline{H}) = 0.5$ for every half-probability event H; in which case $\succsim$ maximizes expected utility relative to p and

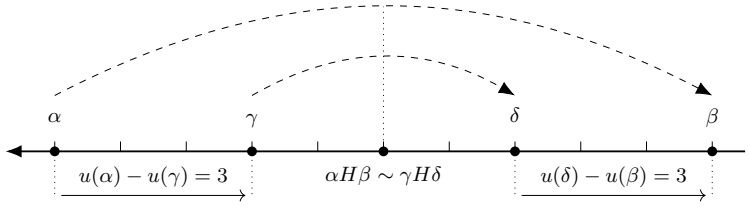

Fig. 3 Employing half-probability events to identify equal utility differences: $\alpha H \beta \sim \gamma H \delta$ implies that the utility interval between α and β shares a midpoint with the utility interval between γ and δ.

some utility function u only if

$$\alpha H \beta \sim \gamma H \delta \;\Leftrightarrow\; \frac{u(\alpha) + u(\beta)}{2} = \frac{u(\gamma) + u(\delta)}{2} \;\Leftrightarrow\; u(\alpha) - u(\gamma) = u(\delta) - u(\beta).$$

In terms of the final representation, $\alpha H \beta \sim \gamma H \delta$ can be thought of as saying that the midway point in utility between α and β is the same as the midway point in utility between γ and δ, from which it follows that the difference in utility between α and γ must be equal to that between δ and β—as represented in Figure 3.

Combining C4 and C5, it can be shown that for any half-probability events H,

$$\alpha H \beta \sim \gamma H \delta \;\Leftrightarrow\; \Delta(\alpha, \gamma) = \Delta(\delta, \beta).$$

This establishes the equal-spacing structure (see Lemma 1). In the presence of that structure, C5 then expresses the fact that relative expected utility depends not on the absolute utility of each consequence at a state, but instead on the differences between those utilities for each state. Compare the following decision matrices, where C5 tells us that $f \succsim g \Leftrightarrow f' \succsim g'$ precisely because the pattern of utility differences $(+2, -1, 0)$ is identical between them:

	s_1	s_2	s_3
f	1	2	4
g	3	1	4

	s_1	s_2	s_3
f'	2	1	3
g'	4	0	3

While C5 is not necessary for consistency with expected utility theory in general, it *is* necessary given an equal-spaced utility distribution. In fact, in that context C5 is a natural strengthening of Savage's postulate P2, which in the finite setting is equivalent to:

P2. $\alpha E f \succsim \alpha E g \;\Leftrightarrow\; \beta E f \succsim \beta E g$

P2 comprises one half of Savage's *Sure Thing Principle*, and follows from C5 by the substitution $(\alpha, \beta, \alpha', \beta') = (\alpha, \alpha, \beta, \beta)$. The other half is given by his postulate P3, which is here relabeled as C6. Say that an event E is *null* just in case $\alpha E f \sim \beta E f$ for all consequences α, β. Then C6 is a simple event-wise dominance condition:

C6. If E is non-null, then $\alpha \succsim \beta \;\Leftrightarrow\; \alpha E f \succsim \beta E f$

C1–C6 suffice for the existence of a utility function u such that

1. $\alpha \succsim \beta \Leftrightarrow u(\alpha) \geq u(\beta)$,
2. $\alpha H \beta \succsim \gamma H \delta \Leftrightarrow u(\alpha) - u(\gamma) \geq u(\delta) - u(\beta)$.

Furthermore, we can show that any utility function satisfying these properties will be unique up to a positive affine transformation. (See Lemma 2.) As Ramsey [30] pointed out, we can then use the fact that difference ratios are always preserved under positive affine transformations to extract a unique normalized confidence function out of u—at least provided certain further conditions obtain.

There's more than one way to do this. Ramsey's original strategy requires, first of all, that every binary act $\alpha E \beta$ has a *certainty equivalent*—i.e., a consequence γ such that $\gamma \sim \alpha E \beta$. This is used to pin the value of every binary act to a particular point on the utility scale. More importantly, the strategy requires presupposing that c satisfies *normalized complement-additivity*—i.e., $c(E) + c(\overline{E}) = 1$ for all E. Given that, $V_{c,u}$ represents $\succsim$ only if

$$c(E) = \frac{u(\gamma) - u(\beta)}{u(\alpha) - u(\beta)},$$

whenever $\alpha \succ \beta$ and $\gamma \sim \alpha E \beta$. Thus Ramsey suggested that we might define $c(E)$ using ratios of differences in the utilities of the consequences of binary acts conditional on E and their certainty equivalents, under the assumption that c satisfies normalized complement-additivity.

For present purposes, it is better if we do not presuppose normalized complement-additivity—we want the structure of c to be a consequence of the assumption that it relates to $\succsim$ via the expected utility rule. Thus we pursue an alternative, more general method that requires neither certainty equivalents nor the assumption of normalized complement-additivity. Instead, the current strategy uses 'exchange rates' between utilities to derive confidence ratios between events. For a simple example of what I mean, consider the following matrix:

	E	$\overline{E}$
f	1	0
g	0	3

Now suppose $f \sim g$. We might think of this intuitively as saying that the prospect of one unit of utility at E can be 'exchanged' for the prospect of 3 units of utility at $\overline{E}$; and in this case, $V_{c,u}$ represents $\succsim$ only if $c(E) = 3 \cdot c(\overline{E})$. More generally, suppose that E and F are non-null events and that there are consequences α, β, γ such that $\alpha, \beta \succ \gamma$ and $\alpha E \gamma \sim \beta F \gamma$. Then $V_{c,u}$ represents $\succsim$ only if

$$\frac{c(E)}{c(F)} = \frac{u(\beta) - u(\gamma)}{u(\alpha) - u(\gamma)}.$$

For this to work, two more conditions are needed. First, we need to ensure coherence in the exchange rates. To that end, observe that the ratio between any two non-null E and E' can be established indirectly via their respective relationships with a common benchmark F:

$$\left.\begin{array}{l} c(F) = \frac{\mathrm{d}}{\mathrm{a}} \cdot c(E) \\ c(F) = \frac{\mathrm{d}}{\mathrm{b}} \cdot c(E') \end{array}\right\} \Rightarrow \mathrm{b} \cdot c(E) = \mathrm{a} \cdot c(E').$$

Thus we have the following condition, which holds necessarily for any preferences that have an expected utility maximizing representation:

C7. If there are consequences $\alpha, \beta, \gamma, \delta$ and events E, E', F such that $\alpha, \beta, \gamma \succ \delta$, $\gamma E \delta \sim \alpha F \delta$, and $\gamma E' \delta \sim \beta F \delta$, then $\beta E \delta \sim \alpha E' \delta$

Finally, there must be sufficiently many consequences of differing values to fix an exchange rate between any two non-null events. Since $\mathcal{S}$ is finite, the following suffices:

C8. There is a non-null state s such that for all non-null states s', there are $\alpha, \beta, \gamma \in \mathcal{C}$ such that $\Delta(\beta, \gamma) \geq \Delta(\alpha, \gamma) = 1$ and $\alpha s' \gamma \sim \beta s \gamma$

The main effect of C8 is to impose a certain kind of factorizability constraint on c: all confidence weights must be a non-negative integer multiple of the weight assigned to some minimal non-null state. This is a strong constraint on the shape of c; however, it also greatly simplifies the existence proof while leaving plenty of room for ordinal underdetermination.

We can now state the main result on which the argument for Preference-Grounded Ratios is premised:

Theorem 2 (PGR theorem) *Suppose that $\mathcal{S}$ and $\mathcal{C}$ are non-empty finite sets, $\succsim$ is a binary relation on the set of finitely-valued functions from $\mathcal{S}$ to $\mathcal{C}$, $\succsim_c$ is a binary relation on $2^{\mathcal{S}}$, and C1–C8 hold. Then there exists a probability function p and a utility function u such that:*

PGR1. $V_{p,u}$ represents $\succsim$,

PGR2. p represents $\succsim_c$.

Furthermore,

PGR3. p is the unique probability function for which PGR1 holds, while u is unique up to a positive affine transformation.

See Appendix A for a proof.

For later reference, it's worth noting that the conditions necessary for the existential conclusions PGR1 and PGR2 are C1, C2, C6 and C7. More explicitly: those conditions hold any time that $\succsim$ maximizes $V_{p,u}$ relative to some p and u, and p represents $\succsim_c$. The remaining conditions all play an important role in either the existence proof or the uniqueness proof, or both, but none of them are required for any of the theorem's conclusions whether taken individually or in combination. That is, it is possible to devise cases where some (or all) of C3, C4, C5 and C8 are violated, and yet each of PGR1, PGR2 and PGR3 still hold.

5 Meaning and ratios

The first step towards establishing Preference-Grounded Ratios is to show that whenever $\succsim$ and $\succsim_c$ jointly satisfy C1–C8, then the pattern of ratios encoded by the unique probability function relative to which $\succsim$ maximizes expected utility is meaningful. Informally—I'll provide a precise characterization just below—a property P of a confidence function is meaningful according to a theory whenever that theory leaves no freedom of choice with respect to whether P is satisfied; in that sense, then, P is not merely an artifact of some arbitrary choice of scale. Or, even more informally, meaningful properties are those about which the theory is telling us: *this corresponds to something real in what's being represented*.

The immediate issue to address is that the PGR theorem, as stated above, does not directly specify which properties of the probabilistic representation are meaningful. This is something it shares with typical statements of Savage's theorem, and of representation theorems for expected utility theory more generally. As Mike Titelbaum put it in a recent discussion,

> The phrase "there exists a unique probabilistic credence distribution" contains a key ambiguity. One might be tempted to read it as saying that given an agent's full preference ranking of acts, there will be exactly one credence distribution that matches those preferences via expected utility maximization, and *moreover* that credence distribution will be probabilistic. But that's not how the theorem works. The proof begins by assuming that we're looking for a probabilistic credence distribution, and then showing that out of all the probabilistic distributions, there is exactly one that will match the agent's preferences... (Titelbaum [25], 300)[11]

While it's not really an ambiguity, it is generally true that the question of whether $\succsim$ might also have a V-maximizing representation relative to some other *non-probabilistic* confidence function—and if so, then how all these functions might relate to one another—is not *explicitly* addressed in the statement of the theorem. Consequently, it is not immediately obvious from such statements which properties of the probabilistic representation are purely the result of a conventional choice to focus specifically on *probabilistic* representations.

In what follows, I will start with a characterization of meaningfulness appropriate for the present context. The definition is designed to fit naturally with standard measurement-theoretic notions of meaningfulness in the numerical representation of quantities—i.e., invariance under automorphisms of a fixed representational structure, as specified by some formal theory of the quantity in question.[12] Once that's in place, it is not difficult to isolate those properties of a confidence function that are meaningful with respect to V-maximizing representations of a preference relation—and thereby show that any proof that $\succsim$ maximizes expected utility relative to a unique probability function p is also, in effect, a proof that the total pattern of ratios encoded by p is meaningful.

---

[11] Similar points are raised by Meacham and Weisberg [24], 657 and Hájek [23], 803. It is worth noting that Titelbaum does go on to argue that ratio information is meaningful in the context of Savage's representation, though the route he takes to get there is a little different from my own.

[12] See, e.g., Pfanzagl [36], 34–35; Luce [37], 3; Narens [38], 155–158; Roberts [39], 58–59; Falmagne & Roberts [40]; Luce et al. [41], 269–282; Mundy [42]. For a detailed treatment of meaningfulness specifically in relation to decision theory, see Elliott [19].

We set aside the case where $\succsim$ is trivial. (If $\succsim$ is trivial, then $V_{c,u}$ represents $\succsim$ whenever u takes constant value 0 *regardless* of what confidence function we plug into the valuation. This tells us nothing interesting about what is meaningful, only that we ought to ignore the degenerate case.) Given that, where P might be any property of a confidence function:

- P is *meaningful* just in case, for every non-trivial $\succsim$ with a V-maximizing representation, any two such representations $V_{c,u}$ and $V_{c',u'}$ agree with respect to P; i.e., c satisfies P if and only if c' satisfies P.

Notice that meaningfulness splits into two kinds. P is *trivially* meaningful whenever there's no variation, across all possible V-maximizing representations of any non-trivial $\succsim$ relation, with respect to whether c satisfies P. In that sense, the question of whether c satisfies P is settled just by the hypothesis that it figures in some V-maximizing representation or other. On the other hand, P is *substantively* meaningful whenever it's meaningful but not trivially meaningful. In effect, every substantively meaningful property marks a distinction within the class of $\succsim$ relations that have some V-maximizing representation: those for which the structure of $\succsim$ forces P, and those for which it instead forces $\neg$P.[13]

It is straightforward to prove that additivity and non-constancy are trivially meaningful, each falling directly out of V-maximization in all but the degenerate case where $\succsim$ is trivial:

Proof Assume that $\succsim$ is non-trivial, and that $V_{c,u}$ represents $\succsim$. From non-triviality, u must be non-constant. Given that, consider any two consequences α and β where $u(\alpha) \neq u(\beta)$, and any pair of disjoint events E and F. Since $\alpha E \cup F \alpha = \alpha E \alpha F \alpha$ and $\beta E \cup F \beta = \beta E \beta F \beta$, the partition-independence requirement on the definition of $V_{c,u}$ (Section 2) implies

$$c(E \cup F) \cdot u(\alpha) + c(\overline{E \cup F}) \cdot u(\alpha) = c(E) \cdot u(\alpha) + c(F) \cdot u(\alpha) + c(\overline{E \cup F}) \cdot u(\alpha),$$
$$c(E \cup F) \cdot u(\beta) + c(\overline{E \cup F}) \cdot u(\beta) = c(E) \cdot u(\beta) + c(F) \cdot u(\beta) + c(\overline{E \cup F}) \cdot u(\beta).$$

Since $u(\alpha) \neq u(\beta)$, together these imply that c is additive: $c(E \cup F) = c(E) + c(F)$. Additivity then implies that $c(\varnothing) = 0$, so c is constant only if $c(E) = 0$ for all E. But this would imply that $\succsim$ is trivial; so c must be non-constant. □

By contrast, neither normalization nor non-negativity are trivially meaningful, since we can readily concoct preferences with a V-maximizing representation whose confidence function satisfies neither property. Suppose that $E, \overline{E}$ are non-empty, and let

$$\begin{array}{lll} c(\varnothing) = 0 & c(E) = 1 & u(\alpha) = 1 \\ c(\mathcal{S}) = 0 & c(\overline{E}) = -1 & u(\beta) = 0 \end{array}$$

[13] Strictly speaking, the definition I've provided specifies when P is (trivially/substantively) meaningful *with respect to V-maximizing representations of preference*. Parallel definitions could be provided for when P is meaningful *with respect to real-valued representations of comparative confidence*. On that basis, we might also say that P is meaningful *simpliciter* whenever it's meaningful with respect to either V-maximizing representations of $\succsim$ *or* real-valued representations of $\succsim_c$. Incorporating this into the discussion would not affect the main conclusion of the section—i.e., that probability ratios are meaningful in the cases where C1–C8 are satisfied—and so I have set it aside for simplicity.

The $V_{c,u}$-induced preference relation therefore includes the following:

$$\alpha E\beta \succ \alpha E\alpha \sim \beta E\beta \succ \beta E\alpha$$

So normalization and non-negativity are not trivially meaningful. Indeed, it turns out that they are not meaningful in either sense. However, observe that those preferences are inconsistent with C6: E is non-null and $\alpha \sim \beta$, yet $\alpha E\beta \succ \beta E\beta$ and $\beta \overline{E}\alpha \succ \alpha \overline{E}\alpha$. This is no accident. Say that c is *zero-bounded* just in case it's either non-negative or non-positive; otherwise it is *mixed-sign*. Then zero-boundedness and mixed-signedness are substantively meaningful properties corresponding precisely to the division, among non-trivial preferences that have at least one V-maximizing representation, of those that do and do not satisfy C6 respectively:

Proof Assume that $\succsim$ is non-trivial, and that $V_{c,u}$ represents $\succsim$. From above, we know that c and u are non-constant, and c is additive, and there is some E with $c(E) \neq 0$. For any such E and any α, β and f,

$$V_{c,u}(\alpha E f) - V_{c,u}(\beta E f) = c(E)\big(u(\alpha) - u(\beta)\big),$$
$$V_{c,u}(\alpha) - V_{c,u}(\beta) = c(\mathcal{S})\big(u(\alpha) - u(\beta)\big).$$

If c is zero-bounded, $c(E)$ and $c(\mathcal{S})$ have the same sign; so

$$V_{c,u}(\alpha) - V_{c,u}(\beta) \geq 0 \Leftrightarrow V_{c,u}(\alpha E f) - V_{c,u}(\beta E f) \geq 0,$$

which implies C6. For the other direction, if c is mixed-sign then there are three possibilities:

1. $c(\mathcal{S}) = 0$,
2. $c(\mathcal{S}) > 0$, and
3. $c(\mathcal{S}) < 0$.

For case 1, pick any α, β such that $u(\alpha) > u(\beta)$, and pick an event E such that $c(E) < 0$ and $c(\overline{E}) > 0$. Observe that $V_{c,u}(\alpha E\alpha) = c(\mathcal{S}) \cdot u(\alpha) = 0$ and likewise $V_{c,u}(\beta E\beta) = c(\mathcal{S}) \cdot u(\beta) = 0$, so $\alpha \sim \beta$. However, we also get that $V_{c,u}(\beta E\alpha) > V_{c,u}(\alpha E\alpha)$ and so $\beta E\alpha \succ \alpha E\alpha$, which violates C6. The proof that cases 2 and 3 also imply a violation of C6 is essentially the same, and therefore omitted. So, c is zero-bounded if and only if $\succsim$ satisfies C6. □

The upshot is that we can now specify the exact role that normalization and non-negativity play with respect to V-maximization—essentially, given additivity, they pin down one particular way of instantiating non-constancy and zero-boundedness. But that is the limit of what is meaningful in them. That is, once non-constancy and zero-boundedness are established, normalization and non-negativity amount to no more than an arbitrary choice of unit and sign respectively. This is implied by the fact that if $V_{c,u}$ represents $\succsim$, and c' is any scalar transformation of c (i.e., $c' = \mathrm{a} \cdot c$ for $\mathrm{a} \neq 0$), then there is a u' such that $V_{c',u'}$ also represents $\succsim$:

Proof Assume that $V_{c,u}$ represents $\succsim$. For any $\mathrm{a} > 0$, let $c' = \mathrm{a} \cdot c$, $c'' = -\mathrm{a} \cdot c$, and $u' = -1 \cdot u$. Then for any f,

$$\mathrm{a} \cdot \sum_{E \in \mathcal{E}_f} c(E) \cdot u(f_E) = \sum_{E \in \mathcal{E}_f} c'(E) \cdot u(f_E) = \sum_{E \in \mathcal{E}_f} c''(E) \cdot u'(f_E).$$

So $\mathrm{a} \cdot V_{c,u}(f) = V_{c',u}(f) = V_{c'',u'}(f)$. Since multiplication by a positive constant is order-preserving, therefore $V_{c',u}$ and $V_{c'',u'}$ also represent $\succsim$. □

There are two lessons to draw from all this. First: $\succsim$ maximizes V relative to c and u *only if* c is additive, non-constant, and unique *at most* up to a scalar transformation. The 'at most' is important—for all I've said, it may yet be the case that $V_{c,u}$ and $V_{c',u'}$ both represent $\succsim$ even if c' is *not* a scalar transformation of c. More generally, it simply isn't possible to fix a uniqueness condition on the confidence function *stronger than* multiplication by a non-zero constant unless further constraints are added on the form of the intended representation, and indeed there will be conditions under which the relevant uniqueness properties are strictly weaker. But now the second lesson: to show that (a) $\succsim$ maximizes expected utility relative to exactly one probability function *just is* to show that (b) $\succsim$ maximizes V relative to an additive, non-constant and zero-bounded confidence function that's unique up to a scalar transformation. The equivalence proof is straightforward, given the points established above:

Proof That (b) implies (a) is immediate. For the other direction, assume that $\succsim$ maximizes expected utility relative to a unique probability function p. $\succsim$ therefore maximizes V relative to every c that's a scalar transformation of p, and *doesn't* maximize V relative to any c that's non-additive, constant or mixed-sign. As such, let $c^\star$ be any additive, non-constant and zero-bounded confidence function that is not a scalar transformation of p, and suppose there's some u such that $V_{c^\star,u}$ represents $\succsim$. Then there must be some $p^\star \neq p$ such that $p^\star$ is a scalar transformation of $c^\star$ and $\succsim$ maximizes expected utility relative to $p^\star$, contradicting the initial assumption. □

We can now address the challenge with which this section began. The PGR theorem asserts that if C1–C8 are satisfied, then there's a unique probability function p and some utility function u such that $V_{p,u}$ represents $\succsim$. It is then a corollary that there exists some u' such that $V_{c,u'}$ represents $\succsim$ if and only if c is a scalar transformation of p. Scalar transformations are precisely those that preserve ratio information, in the sense that c is a scalar transformation of c' if and only if c and c' agree on which events receive value 0, and for all E, F such that $c(F) \neq 0$,

$$\frac{c(E)}{c(F)} = \frac{c'(E)}{c'(F)}.$$

Conclusion: if C1–C8 are satisfied, then the structure of the preference relation forces a *total pattern of ratios among non-null events* in any confidence function to which those preferences relate via V-maximization. From the perspective of expected utility theory, then, there is no meaningful difference between two confidence functions which instantiate that pattern.

6 When comparative confidence isn't enough

The next step is to argue that there exist at least some circumstances under which C1–C8 are satisfied, and yet the resulting ratios cannot be explained wholly in terms of $\succsim_c$. For this it suffices to find a single instance involving an ordinally underdetermined probability function. With that said, the structure of the argument is a little more transparent if we instead find a *pair* of cases involving two ordinally underdetermined probability functions that represent the same confidence ordering.

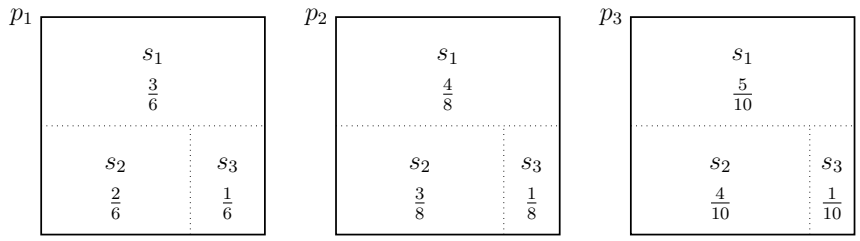

Fig. 4 Eikosograms for three probability functions all representing the same confidence ordering.

To that end, consider the probability functions p_1, p_2 and p_3 represented in Figure 4, which all represent the following confidence ordering:

$$\mathcal{S} \succ_c E_{12} \succ_c E_{13} \succ_c E_{23} \sim_c s_1 \succ_c s_2 \succ_c s_3 \succ_c \varnothing$$

Set aside p_3 for now; it will be relevant again a little later on. Given $\succsim_c$ as just specified, first we show that there are preferences $\succsim_1$ and $\succsim_2$ such that

1. $\succsim_1$ and $\succsim_c$ satisfy C1–C8, with p_1 being the unique probability function relative to which $\succsim_1$ maximizes expected utility; and
2. $\succsim_2$ and $\succsim_c$ satisfy C1–C8, with p_2 being the unique probability function relative to which $\succsim_2$ maximizes expected utility.

The simplest possible example requires five consequences. Let $\mathcal{C} = \{0, 1, 2, 3, 4\}$, and set $u(i) = i$ for $i = 0, \ldots, 4$. In that case, Table 1 represents the ordering $\succsim_1$ induced by $V_{p_1,u}$, while Table 2 represents the ordering $\succsim_2$ induced by $V_{p_2,u}$. Moreover, it is easy to see that both examples satisfy C1–C8. C3 follows from $E_{23} \sim_c s_1$. The necessary conditions C1, C2, C6 and C7 follow directly from the fact that the acts in each table are arranged by expected utility. Furthermore, since the utility function is equal-spaced by hypothesis, it is immediate that C4 and C5 are satisfied. Finally, C8 simply requires the existence of a non-null state—in both cases it's s_3—such that for some $x, y > 0$, $f \sim g$ and $f' \sim g'$:

	s_1	s_2	s_3
f	1	0	0
g	0	0	x

	s_1	s_2	s_3
f'	0	1	0
g'	0	0	y

For p_1, $x = 3$ and $y = 2$; and for p_2, $x = 4$ and $y = 3$.

So p_1 is the unique probability function relative to which $\succsim_1$ maximizes expected utility, while p_2 is the unique probability function relative to which $\succsim_2$ maximizes expected utility, even though p_1 and p_2 represent the same confidence ordering. Since we have already concluded that, under these circumstances, the ratio relations encoded within p_1 and p_2 are meaningful, the upshot is that there must be more to what makes them meaningful than can be explained solely in terms of $\succsim_c$. More directly, what makes them meaningful in this case are the exchange rates—which are fundamentally a matter of the relationship between confidence, utility and preference.

s_1	s_2	s_3	eu	s_1	s_2	s_3	eu	s_1	s_2	s_3	eu	s_1	s_2	s_3	eu
0	0	0	0.0	0	2	4	1.33	2	3	0	2.0	3	2	3	2.67
0	0	1	0.17	1	2	1	1.33	3	0	3	2.0	3	3	1	2.67
0	0	2	0.33	0	3	3	1.5	3	1	1	2.0	4	0	4	2.67
0	1	0	0.33	0	4	1	1.5	4	0	0	2.0	4	2	0	2.67
0	0	3	0.5	1	1	4	1.5	1	4	2	2.17	2	4	3	2.83
0	1	1	0.5	1	2	2	1.5	3	1	2	2.17	3	3	2	2.83
1	0	0	0.5	1	3	0	1.5	1	3	4	2.17	3	4	0	2.83
0	1	2	0.67	2	0	3	1.5	2	2	3	2.17	4	1	3	2.83
0	0	4	0.67	2	1	1	1.5	2	3	1	2.17	3	2	4	2.83
0	2	0	0.67	3	0	0	1.5	3	0	4	2.17	4	2	1	2.83
1	0	1	0.67	0	4	2	1.67	3	2	0	2.17	2	4	4	3.0
0	1	3	0.83	2	1	2	1.67	4	0	1	2.17	3	3	3	3.0
1	0	2	0.83	0	3	4	1.67	1	4	3	2.33	3	4	1	3.0
1	1	0	0.83	1	2	3	1.67	2	3	2	2.33	4	1	4	3.0
0	2	1	0.83	1	3	1	1.67	2	4	0	2.33	4	2	2	3.0
0	1	4	1.0	2	0	4	1.67	3	1	3	2.33	4	3	0	3.0
0	2	2	1.0	2	2	0	1.67	4	0	2	2.33	3	4	2	3.17
0	3	0	1.0	3	0	1	1.67	4	1	0	2.33	3	3	4	3.17
1	0	3	1.0	0	4	3	1.83	2	2	4	2.33	4	2	3	3.17
1	1	1	1.0	1	3	2	1.83	3	2	1	2.33	4	3	1	3.17
2	0	0	1.0	1	4	0	1.83	1	4	4	2.5	3	4	3	3.33
1	1	2	1.17	2	1	3	1.83	2	3	3	2.5	4	3	2	3.33
0	2	3	1.17	3	0	2	1.83	2	4	1	2.5	4	4	0	3.33
0	3	1	1.17	3	1	0	1.83	3	1	4	2.5	4	2	4	3.33
1	0	4	1.17	1	2	4	1.83	3	2	2	2.5	3	4	4	3.5
1	2	0	1.17	2	2	1	1.83	3	3	0	2.5	4	3	3	3.5
2	0	1	1.17	0	4	4	2.0	4	0	3	2.5	4	4	1	3.5
0	3	2	1.33	1	3	3	2.0	4	1	1	2.5	4	4	2	3.67
0	4	0	1.33	1	4	1	2.0	2	4	2	2.67	4	3	4	3.67
1	1	3	1.33	2	1	4	2.0	4	1	2	2.67	4	4	3	3.83
2	0	2	1.33	2	2	2	2.0	2	3	4	2.67	4	4	4	4.0
2	1	0	1.33												

Table 1 Preferences $\succsim_1$, induced by $u(i) = i$ and $p(s_1) = \frac{3}{6}$, $p(s_2) = \frac{2}{6}$, $p(s_3) = \frac{1}{6}$.

s_1	s_2	s_3	eu	s_1	s_2	s_3	eu	s_1	s_2	s_3	eu	s_1	s_2	s_3	eu
0	0	0	0.0	1	1	4	1.38	3	0	4	2.0	4	1	2	2.63
0	0	1	0.13	2	0	3	1.38	3	1	1	2.0	3	3	1	2.75
0	0	2	0.25	2	1	0	1.38	4	0	0	2.0	2	4	2	2.75
0	0	3	0.38	0	3	3	1.5	1	3	4	2.13	3	2	4	2.75
0	1	0	0.38	0	4	0	1.5	1	4	1	2.13	4	2	0	2.75
0	0	4	0.5	1	2	2	1.5	2	3	0	2.13	4	1	3	2.75
0	1	1	0.5	2	0	4	1.5	4	0	1	2.13	3	3	2	2.88
1	0	0	0.5	2	1	1	1.5	2	2	3	2.13	4	2	1	2.88
1	0	1	0.63	3	0	0	1.5	3	1	2	2.13	2	4	3	2.88
0	1	2	0.63	0	3	4	1.63	2	3	1	2.25	4	1	4	2.88
0	2	0	0.75	0	4	1	1.63	1	4	2	2.25	2	4	4	3.0
1	0	2	0.75	1	3	0	1.63	2	2	4	2.25	3	3	3	3.0
0	1	3	0.75	3	0	1	1.63	3	2	0	2.25	3	4	0	3.0
0	2	1	0.88	1	2	3	1.63	4	0	2	2.25	4	2	2	3.0
0	1	4	0.88	2	1	2	1.63	3	1	3	2.25	3	3	4	3.13
1	0	3	0.88	1	3	1	1.75	2	3	2	2.38	3	4	1	3.13
1	1	0	0.88	0	4	2	1.75	3	2	1	2.38	4	3	0	3.13
0	2	2	1.0	1	2	4	1.75	1	4	3	2.38	4	2	3	3.13
1	0	4	1.0	2	2	0	1.75	3	1	4	2.38	4	3	1	3.25
1	1	1	1.0	3	0	2	1.75	4	0	3	2.38	3	4	2	3.25
2	0	0	1.0	2	1	3	1.75	4	1	0	2.38	4	2	4	3.25
0	3	0	1.13	1	3	2	1.88	1	4	4	2.5	4	3	2	3.38
2	0	1	1.13	2	2	1	1.88	2	3	3	2.5	3	4	3	3.38
0	2	3	1.13	0	4	3	1.88	2	4	0	2.5	3	4	4	3.5
1	1	2	1.13	2	1	4	1.88	3	2	2	2.5	4	3	3	3.5
0	3	1	1.25	3	0	3	1.88	4	0	4	2.5	4	4	0	3.5
0	2	4	1.25	3	1	0	1.88	4	1	1	2.5	4	3	4	3.63
1	2	0	1.25	0	4	4	2.0	2	3	4	2.63	4	4	1	3.63
2	0	2	1.25	1	3	3	2.0	2	4	1	2.63	4	4	2	3.75
1	1	3	1.25	1	4	0	2.0	3	3	0	2.63	4	4	3	3.88
0	3	2	1.38	2	2	2	2.0	3	2	3	2.63	4	4	4	4.0
1	2	1	1.38												

Table 2 Preferences $\succsim_2$, induced by $u(i) = i$ and $p(s_1) = \frac{4}{8}$, $p(s_2) = \frac{3}{8}$, $p(s_3) = \frac{1}{8}$.

E	F	$p_1(E)$	$p_1(F)$	$p_1(E\|F)$
E_{12}	$\mathcal{S}$	5/6	6/6	$5/6 \approx 0.83$
s_1	E_{13}	3/6	4/6	$3/4 = 0.75$
s_2	E_{23}	2/6	3/6	$2/3 \approx 0.67$
E_{13}	$\mathcal{S}$	4/6	6/6	$4/6 \approx 0.67$
s_1	E_{12}	3/6	5/6	$3/5 = 0.60$
s_1	$\mathcal{S}$	3/6	6/6	$3/6 = 0.50$
E_{23}	$\mathcal{S}$	3/6	6/6	$3/6 = 0.50$
s_2	E_{12}	2/6	5/6	$2/5 = 0.40$
s_2	$\mathcal{S}$	2/6	6/6	$2/6 \approx 0.33$
s_3	E_{23}	1/6	3/6	$1/3 \approx 0.33$
s_3	E_{13}	1/6	4/6	$1/4 = 0.25$
s_3	$\mathcal{S}$	1/6	6/6	$1/6 \approx 0.17$

E	F	$p_2(E)$	$p_2(F)$	$p_2(E\|F)$
E_{12}	$\mathcal{S}$	7/8	8/8	$7/8 \approx 0.88$
s_1	E_{13}	4/8	5/8	$4/5 = 0.80$
s_2	E_{23}	3/8	4/8	$3/4 = 0.75$
E_{13}	$\mathcal{S}$	5/8	8/8	$5/8 \approx 0.63$
s_1	E_{12}	4/8	7/8	$4/7 \approx 0.57$
s_1	$\mathcal{S}$	4/8	8/8	$4/8 = 0.50$
E_{23}	$\mathcal{S}$	4/8	8/8	$4/8 = 0.50$
s_2	E_{12}	3/8	7/8	$3/7 \approx 0.43$
s_2	$\mathcal{S}$	3/8	8/8	$3/8 \approx 0.38$
s_3	E_{23}	1/8	4/8	$1/4 = 0.25$
s_3	E_{13}	1/8	5/8	$1/5 = 0.20$
s_3	$\mathcal{S}$	1/8	8/8	$1/8 \approx 0.13$

Table 3 Comparative confidence orders induced by p_1 (left) and p_2 (right). Notes: conditional probabilities $p(E \,|\, F)$ are represented only for the cases where $E \subset F$ and $E \neq \varnothing$. The trivial cases where $E = \varnothing$ (so $p(E|F) = 0$) and $E \cap F = F$ (so $p(E|F) = 1$) have been excluded for simplicity, and it is assumed that $\succsim_k$ compares all and only pairs whose conditioning events are non-null. (In this case, the only null event is $\varnothing$.) Given that, all other $\succsim_k$-comparisons—i.e., those where $E \not\subseteq F$ and $E \cap F \neq \varnothing$—are fully determined by those comparisons presented in the table.

But one might be worried that this argument depends on a relatively impoverished understanding of comparative confidence—in particular, rather than focusing on an *unconditional* confidence relation over individual events, perhaps we should have instead been using *conditional* confidence relations.[14] Let $\succsim_k$ be a binary relation over $2^{\mathcal{S}} \times 2^{\mathcal{S}}$, where $(E, E') \succsim_k (F, F')$ is taken to mean that the agent has at least as much confidence in E *given* E' as in F *given* F'. In that case, we can treat *unconditional* confidence comparisons as just comparisons conditional on the necessary event; so

$$E \succsim_c F \;\Leftrightarrow\; (E, \mathcal{S}) \succsim_k (F, \mathcal{S}).$$

Moreover, since $\succsim_k$ includes comparisons not captured by $\succsim_c$, it permits strictly more fine-grained differentiation among probability functions. Most notably, even though p_1 and p_2 represent the same *unconditional* confidence ordering, they correspond to very different *conditional* confidences (see Table 3).

Does this mean that the richer structure of these conditional confidence relations will be enough to determine a unique probabilistic representation even in cases where the corresponding unconditional confidence relation cannot? Sometimes. But consider p_2 and p_3, which are equivalent with respect to both $\succsim_c$ and $\succsim_k$ (see Table 4). And just as there are preference orders $\succsim_1$ and $\succsim_2$ that maximize expected utility relative to p_1 and p_2 respectively, likewise it is easy to show that there is a preference order $\succsim_3$ that maximizes expected utility relative to p_3 (and not p_1 or p_2). Consequently, there are cases in which C1–C8 are all satisfied, and yet the resulting probability ratios cannot be explained wholly in terms of $\succsim_c$ *or* in terms of $\succsim_k$. Merely appealing to the richer structure of conditional confidence doesn't yet address the challenge; all it does is move the bump under the rug.

[14] Thanks to a referee for suggesting further discussion on this point. For more in relation to conditional confidence relations, see Koopman [4], [5], Luce [8], Suppes and Zanotti [12], Joyce [28], DiBella [16], and Hawthorne [43].

E	F	$p_2(E)$	$p_2(F)$	$p_2(E\|F)$
E_{12}	$\mathcal{S}$	7/8	8/8	$7/8 \approx 0.88$
s_1	E_{13}	4/8	5/8	$4/5 = 0.80$
s_2	E_{23}	3/8	4/8	$3/4 = 0.75$
E_{13}	$\mathcal{S}$	5/8	8/8	$5/8 \approx 0.63$
s_1	E_{12}	4/8	7/8	$4/7 \approx 0.57$
s_1	$\mathcal{S}$	4/8	8/8	$4/8 = 0.50$
E_{23}	$\mathcal{S}$	4/8	8/8	$4/8 = 0.50$
s_2	E_{12}	3/8	7/8	$3/7 \approx 0.43$
s_2	$\mathcal{S}$	3/8	8/8	$3/8 \approx 0.38$
s_3	E_{23}	1/8	4/8	$1/4 = 0.25$
s_3	E_{13}	1/8	5/8	$1/5 = 0.20$
s_3	$\mathcal{S}$	1/8	8/8	$1/8 \approx 0.13$

E	F	$p_3(E)$	$p_3(F)$	$p_3(E\|F)$
E_{12}	$\mathcal{S}$	9/10	10/10	$9/10 = 0.90$
s_1	E_{13}	5/10	6/10	$5/6 \approx 0.83$
s_2	E_{23}	4/10	5/10	$4/5 = 0.80$
E_{13}	$\mathcal{S}$	6/10	10/10	$3/5 = 0.60$
s_1	E_{12}	5/10	9/10	$5/9 \approx 0.56$
s_1	$\mathcal{S}$	5/10	10/10	$5/10 = 0.50$
E_{23}	$\mathcal{S}$	5/10	10/10	$5/10 = 0.50$
s_2	E_{12}	4/10	9/10	$4/9 \approx 0.44$
s_2	$\mathcal{S}$	4/10	10/10	$2/5 = 0.40$
s_3	E_{23}	1/10	5/10	$1/5 = 0.20$
s_3	E_{13}	1/10	6/10	$1/6 \approx 0.17$
s_3	$\mathcal{S}$	1/10	10/10	$1/10 = 0.10$

Table 4 Comparative confidence orders induced by p_2 (left) and p_3 (right).

One might also be worried that the possibility of these kinds of examples depends on the assumption that $\mathcal{S}$ is finite.[15] There are two responses to this. One: *it almost certainly doesn't*; and two: *it wouldn't matter either way.* But before that, it is important to first note that there can be probability functions defined over infinite state spaces that are underdetermined by the confidence orderings to which they correspond. I will give two simple examples; they are not the only ones. First, assume that $\mathcal{S}$ is infinite, and split it into two infinite subsets $\mathcal{S}_a$, $\mathcal{S}_b$. Let U_a be any (principal or non-principal) ultrafilter on $\mathcal{S}_a$, and define the probability function p_a over the powerset of $\mathcal{S}_a$ such that $p_a(E) = 1$ if $E \in U_a$, otherwise $p_a(E) = 0$.[16] Do the same for $\mathcal{S}_b$, *mutatis mutandis*. Now take any $\mathrm{r} \in (0, \frac{1}{2})$ and define p_r on $\mathcal{S}$ such that

$$p_\mathrm{r}(E) = \mathrm{r} \cdot p_a(E \cap \mathcal{S}_a) + (1 - \mathrm{r}) \cdot p_b(E \cap \mathcal{S}_b).$$

By comparing different choices of r, one can verify that p_r is underdetermined by $\succsim_c$; and indeed, it is also underdetermined by $\succsim_k$. But perhaps you'll think this example doesn't *really* count—if both ultrafilters are principal, then the construction is effectively just a finite-support probability measure padded out with infinitely many null states; and if either is non-principal, then the measure is only finitely additive. So, for the second example, let $\mathcal{S} = \mathbb{N}$, and for each $\mathrm{r} \in (0, \frac{1}{2})$, set

$$p_\mathrm{r}(\{n\}) = (1 - \mathrm{r})\mathrm{r}^{n-1}.$$

In this case, p_r is countably additive and the atomic probabilities are tail-dominating in the sense that every state is assigned more probability than the entire tail after it. As such, the corresponding $\succsim_c$ ordering just depends on the first state on which two events differ, and that ordering is the same for every $\mathrm{r} \in (0, \frac{1}{2})$.[17]

[15] Thanks to a (different) referee for pushing me to discuss this further.

[16] A *filter* U on $\mathcal{S}$ is a set of subsets of $\mathcal{S}$ such that (i) $\mathcal{S} \in U$; (ii) $\varnothing \notin U$, (iii) if $E \in U$ and $E \subseteq F \subseteq \mathcal{S}$, then $F \in U$, and (iv) if $E \in U$ and $F \in U$, then $E \cap F \in U$. An *ultrafilter* is a filter U such that for all $E \subseteq \mathcal{S}$, either $E \in U$ or $\overline{E} \in U$. An ultrafilter U is *principal* when there is an s such that $U = \{E \subseteq \mathcal{S} : s \in E\}$.

[17] This second example shows that an unconditional confidence ordering $\succsim_c$ can sometimes underdetermine the corresponding probabilities even when it's defined over infinitely many non-null states. Note that this example does *not* establish underdetermination by the conditional confidence ordering $\succsim_k$.

The point of the two examples is that the *size* of $\mathcal{S}$ isn't immediately relevant to whether $\succsim_c$ (or $\succsim_k$) admits a unique probabilistic representation—what matters, rather, is how the subsets of $\mathcal{S}$ are ordered.[18] Setting aside the trivial case where $\mathcal{S}$ contains only a single state, for any state space whatsoever there are some confidence orders that have a unique probabilistic representation and some that have many probabilistic representations. The question, therefore, is whether there's something about the 'exchange rates' reasoning described in Section 4—i.e., the route by which probabilities were derived from utilities—which might essentially depend on the finiteness assumption. But there the answer is clearly no. In general, once interval-scale utilities are fixed, every probability function corresponds to a unique system of exchange rates, which can in principle be used to distinguish it from any other probability function. Given that, for *any* ordinally underdetermined probability function p, *if* there are enough consequences of varying utility around, we can exploit those exchange rates to show that some preference relation $\succsim$ maximizes expected utility uniquely relative to p. The assumption that $\mathcal{S}$ is finite merely serves to make certain parts of the construction conceptually easier, but plays no essential role in the argument. So swapping a finite state space for an infinite one doesn't even move the bump under the rug—it just misses the bump entirely.

But though I've said much about the first response, I emphasize more the second. My argument requires the *existence* of a certain kind of example—whether that example is finite or infinite is inconsequential to the conclusion, so long as it *is*. I see no compelling reason to think that confidence is only ever properly defined over an infinite state space; indeed, it is usually considered an important limitation of many representation theorems that they are incompatible with finite state spaces. So unless there is something philosophically suspect about finite examples in general, what exactly is the problem with appealing to such cases in order to help tease apart explanatory possibilities that might otherwise be difficult to differentiate?

7 What's the alternative?

When all else fails, there is always another option: *just posit more structure*. In fact there are a few ways one might go about this. If $\succsim_c$ (or $\succsim_k$) isn't enough on its own, then perhaps we should add even more primitive relations. A common suggestion along these lines is to supplement comparative confidence with a primitive *independence* relation $\perp$, with the requirement that $E \perp F$ just in case $p(E \cap F) = p(E) \cdot p(F)$.[19] Another possibility would be to say that $\succsim_c$ (or $\succsim_k$) is *really* just the shadow of some much richer relation defined over an expanded domain—some refinement of $\mathcal{S}$, perhaps; or some richer algebra generated from $\mathcal{S}$, such as the Suppes and Zanotti algebra of extended indicator functions.

[18] There is a widespread myth among philosophers that achieving a unique probabilistic representation of a confidence ordering is simply a matter of having an infinite space of events. This likely derives from Savage [26], 596, who described his 'partition assumption' (cf. P6/P6′) as 'an assumption of a sufficient richness of contemplated events'. The statement is a little misleading. Savage's assumption guarantees that every event can be partitioned into arbitrarily many equally probable events. This, however, is not an intrinsic property of the event space. Rather, it is a property that an event space has in relation to a particular kind of probability function—and one that must therefore be reflected in a rather strong structural condition on the corresponding confidence ordering.

[19] See Domotor [44], Kaplan and Fine [45], Luce and Narens [46], Joyce [47].

Of course, merely *positing* this additional structure doesn't automatically solve the explanatory problem. If it is to be anything more than a mere mathematical possibility, one also needs to provide a plausible psychological interpretation of the structure, and under that interpretation, show that it will suffice to explain probability ratios in *all* relevant cases. For example, it would not be enough to simply say that for any state space $\mathcal{S}$ there is always some refinement $\mathcal{R}$ of $\mathcal{S}$ such that the agent's comparative confidence ordering is just a proper part of another relation over $2^{\mathcal{R}}$. One also needs to argue that this richer relation corresponds to something psychologically real and will, as a matter of necessity, have the kind of structure required to pin down at least the ratios between all non-null events in $2^{\mathcal{S}}$. That is *far* from trivial—and without it, the bump remains under the rug.

But I leave that open as a theoretical possibility for others to pursue. For myself, the question isn't whether positing more structure could in principle lead to other explanatory options—of course it can. Rather, the question is whether there's a compelling reason to posit this additional structure in the first place. Unless it's independently well-motivated, the additional structure is only being introduced to help explain ratio information that *already* has a perfectly adequate explanation in terms of how confidence and preference relate to one another in decisions under uncertainty. That explanation can be captured in terms of exchange rates: double the confidence weight at a state and you double the contribution of the utilities of the consequences at that state to the expected value of the act. The ratio inherits its meaning not from any structure internal to confidence itself—or not only—but from the role that confidence plays in relation to other psychological quantities.

It is worth keeping in mind that this kind of explanation is not uncommon. We can in general distinguish between two ways of explaining ratio information in the theory of measurement—they don't have standard names, so let's call them *independent* and *interactive*. Some ratio-scale quantities have independent explanations: it is possible to explain what makes their ratios meaningful *without essential reference to any other quantities*. Size (length, area, volume) is like this in general, and the explanation boils down to the same kind of 'equal-parts' reasoning described in Section 3—e.g., a is n times bigger than b if and only if n duplicates of b are the same size as a. In this case, one doesn't need to know anything about the weights of a and b, for example, or their temperature, surface reflectances, and so on, in order to fully understand what it means for a to be $\frac{\mathrm{n}}{\mathrm{m}}$ times the size of b.

For other quantities, however, the situation is more complicated. For example, one cannot characterize the relational structures that underwrite the ratio-scale measurement of density without some reference to how density relates to other quantities. If a is n times denser than b, then combining n duplicates of b won't result in an object with the same density as a—not unless some mass is added or some volume is lost in the process. Instead, where mass (m) and volume (V) are represented in the usual way, the meaningfulness of density ratios can be explained by the fact that

$$\frac{\rho_i}{\rho_j} = \frac{m_i}{m_j} \cdot \frac{V_j}{V_i},$$

where the ratios on the right-hand side are all scale-independent.

The lesson here is not that mass and volume are *primitive* quantities while density is *derived*, as if this distinction somehow reflects any kind of necessary conceptual or ontological hierarchy between them. As is well known, we could just as well treat density as a primitive while treating mass or volume, or both, as the derived quantities. The difference between primitive/derived in this context is not metaphysical; it just corresponds to the order in which we choose to derive certain quantities from others in a dimensionally coherent system of units. Rather, the lesson is that the three quantities together comprise part of a *system*, and understanding how the parts of that system relate to one another is key to explaining what makes density ratios *meaningful*.

Along the same lines: if you want to know what's meaningful in probabilistic representations of confidence, then one place to look is the role they play in decision theory. That need not be the whole story, but there are good reasons to think it's at least *part* of the story. What value is there in denying ourselves these explanatory resources?

Appendix A Proof of the PGR theorem

The proof of Theorem 2 proceeds as follows. First we use C1–C5 to show that binary acts on half-probability events behave as expected given the intended representation (Lemma 1). Along with C6, this lets us construct an equal-spaced utility function that represents $\succsim$ over $\mathcal{C}$ and is unique up to a positive affine transformation (Lemma 2). We then add C7 and C8 to show that $\succsim$ maximizes expected utility relative to at least one probability function (Lemma 3). Given the facts about expected utility maximization established in Section 5, the PGR theorem then follows readily.

Lemma 1 *If C1–C5 are satisfied, then for all consequences $\alpha, \beta, \gamma, \delta$ and all half-probability events H, H':*

H1. $\alpha H \beta \sim \beta H' \alpha$

H2. If $\succsim$ is non-trivial then H is non-null

H3. $\Delta(\alpha, \gamma) = \Delta(\delta, \beta) \Leftrightarrow \alpha H \beta \sim \gamma H \delta$

H4. If $\alpha H \beta \succ \gamma H \delta$ then $\Delta(\alpha, \gamma) > \Delta(\delta, \beta)$

Proof H1. Consider the following decision matrix:

	$H \cap H'$	$H \cap \overline{H'}$	$\overline{H} \cap H'$	$\overline{H} \cap \overline{H'}$
$\alpha H \beta$	α	α	β	β
$\alpha H' \beta$	α	β	α	β
$\beta H \alpha$	β	β	α	α
$\beta H' \alpha$	β	α	β	α

Note that, since the consequences of $\alpha H \beta$ and $\alpha H' \beta$ are identical under the event $H \cap H'$, C5 implies $\alpha H \beta \succsim \alpha H' \beta \Leftrightarrow \beta H' \alpha \succsim \beta H \alpha$. By completeness, either $\alpha H \beta \succsim \alpha H' \beta$ or $\alpha H' \beta \succsim \alpha H \beta$. Suppose it's the former, and assume (without loss of generality) that $\alpha \succsim \beta$. We know that $\alpha H \beta \sim \beta H \alpha$ for all half-probability events H, since $H \sim_c \overline{H}$ entails

$\alpha H\beta \sim \alpha\overline{H}\beta = \beta H\alpha$ under C2. As such, C5 implies $\alpha H\beta \succsim \alpha H'\beta \sim \beta H'\alpha \succsim \beta H\alpha$, hence $\alpha H\beta \sim \beta H'\alpha$ follows by C1. If it's the latter, then (for the same reasons) we get $\alpha H\beta \sim \beta H'\alpha$. So either way, $\alpha H\beta \sim \beta H'\alpha$.

H2. Suppose that $\succsim$ is non-trivial, with $\alpha \succ \beta$. By C1, if $\alpha H\beta \sim \alpha$ then $\alpha H\beta \not\sim \beta$, whereas if $\alpha H\beta \sim \beta$ then $\beta H\alpha \not\sim \alpha$. So at least one of $\alpha H\beta \not\sim \alpha$ or $\alpha H\beta \not\sim \beta$ holds, and either way H is non-null.

H3. For the left-to-right direction, suppose that $\Delta(\alpha,\gamma) = \Delta(\delta,\beta)$. Then from C5, $\alpha H\beta \sim \delta H\gamma \Leftrightarrow \gamma H\beta \sim \beta H\gamma$. The right-hand side is guaranteed by H1, so $\Delta(\alpha,\gamma) = \Delta(\delta,\beta)$ implies $\alpha H\beta \sim \delta H\gamma \sim \gamma H\delta$. The right-to-left direction is immediate from C4 plus H1.

H4. Observe that, from C4 and H3, for any consequences α, β, there are (not necessarily distinct) consequences γ, δ such that $\alpha H\beta \sim \gamma H\delta$ and there's no further ϵ where $\gamma \succ \epsilon \succ \delta$. There are two possibilities: either $\Delta(\alpha,\beta)$ is even, in which case there is a $\gamma = \delta$ such that $\Delta(\alpha,\gamma) = \Delta(\gamma,\beta)$ and therefore $\alpha H\beta \sim \gamma H\gamma$; or $\Delta(\alpha,\beta)$ is odd, in which case there will be adjacent γ, δ such that $\Delta(\alpha,\gamma) = \Delta(\delta,\beta)$, and therefore $\alpha H\beta \sim \gamma H\delta$. In either case, there will be the same number of consequences between α and $\alpha H\beta$ as there are between $\alpha H\beta$ and β. Given this, assume $\alpha H\beta \succ \gamma H\delta$. There are three ways this might arise:

1. $\alpha \succ \gamma$ and $\beta \succ \delta$,
2. $\alpha \succ \gamma$, $\delta \succsim \beta$, and $\Delta(\alpha,\gamma) > \Delta(\delta,\beta)$, or
3. $\gamma \succsim \alpha$, $\beta \succ \delta$ and $\Delta(\gamma,\alpha) < \Delta(\beta,\delta)$.

In all three circumstances it follows that $\Delta(\alpha,\gamma) > \Delta(\delta,\beta)$. □

Lemma 2 *If C1–C6 are satisfied, then there is a utility function u such that for all $\alpha, \beta, \gamma, \delta$ and all half-probability events H,*

U1. $\alpha \succsim \beta \Leftrightarrow u(\alpha) \geq u(\beta)$

U2. $\alpha H\beta \succsim \gamma H\delta \Leftrightarrow u(\alpha) - u(\gamma) \geq u(\delta) - u(\beta)$

U3. $\Delta(\alpha,\beta) = \Delta(\gamma,\delta) \Leftrightarrow u(\alpha) - u(\beta) = u(\gamma) - u(\delta)$

U4. If any other utility function u' satisfies U1–U3, then there are constants $\mathrm{a} > 0$ *and* b *such that* $u' = \mathrm{a} \cdot u + \mathrm{b}$

Proof Assume that C1–C6 hold. If $\succsim$ is trivial then the result is immediate, so assume it's non-trivial. Let $\alpha\beta \succsim_h \gamma\delta$ mean that $\alpha H\beta \succsim \gamma H\delta$ for some half-probability event H. Given H1, if $\alpha H\beta \succsim \gamma H\delta$ is true for some half-probability event H then it's true for all, so it doesn't matter which H we choose. We start by showing that $\langle \mathcal{C}\times\mathcal{C}, \succsim_h \rangle$ is an additive conjoint structure. The following definition is modified from Krantz et al. [9], since finiteness obviates the usual need for an Archimedean condition:

Definition 1 Suppose that $\mathbf{X}$ and $\mathbf{Y}$ are non-empty finite sets and that $\succsim$ is a binary relation on $\mathbf{X}\times\mathbf{Y}$. Then $\langle \mathbf{X}\times\mathbf{Y}, \succsim \rangle$ is an *additive conjoint structure* just in case $\succsim$ satisfies the following:

ACS1. $\succsim$ is a weak order

ACS2. For all $x, x', x'' \in \mathbf{X}$ and all $y, y', y'' \in \mathbf{Y}$, $xy'' \sim x''y'$ and $x''y \sim x'y''$ implies $xy \sim x'y'$

ACS3. (a) For all $x, x' \in \mathbf{X}$, $xy \succsim x'y$ for any $y \in \mathbf{Y}$ implies $xy' \succsim x'y'$ for all $y' \in \mathbf{Y}$; and (b) for all $y, y' \in \mathbf{Y}$, $xy \succsim xy'$ for any $x \in \mathbf{X}$ implies $x'y \succsim x'y'$ for all $x' \in \mathbf{X}$

ACS4. (a) There are $x, x' \in \mathbf{X}$ and $y \in \mathbf{Y}$ such that $xy \not\sim x'y$; and (b) there are $y, y' \in \mathbf{Y}$ and $x \in \mathbf{X}$ such that $xy \not\sim xy'$

ACS5. (a) If there are $x, x', x'' \in \mathbf{X}$ and $y, y' \in \mathbf{Y}$ for which $x'y' \succsim xy \succsim x''y'$, then there is an $x''' \in \mathbf{X}$ such that $x'''y' \sim xy$; and (b) if there are $x, x' \in \mathbf{X}$ and $y, y', y'' \in \mathbf{Y}$ for which $x'y' \succsim xy \succsim x'y''$, then there is a $y''' \in \mathbf{Y}$ such that $x'y''' \sim xy$

That $\langle \mathcal{C} \times \mathcal{C}, \succsim_h \rangle$ satisfies ACS1 follows immediately from C1. For ACS2, assume $\alpha H \gamma' \sim \gamma H \beta'$ and $\gamma H \alpha' \sim \beta H \gamma'$. By C4 + H1 these imply $\Delta(\alpha, \beta') = \Delta(\beta, \alpha')$; then from H3 we get $\alpha H \alpha' \sim \beta H \beta'$. ACS3(a) follows immediately from C5. ACS4(a) follows immediately from C6 + H2, given that $\succsim$ is non-trivial. For ACS5(a), assume $\gamma' H \delta \succ \alpha H \beta \succ \gamma'' H \delta$. (If either $\succ$ is replaced by $\sim$ then the conclusion is immediate.) By H4 this implies $\Delta(\gamma', \alpha) > \Delta(\beta, \delta) > \Delta(\gamma'', \alpha)$. There are two cases: (i) $\Delta(\beta, \delta) \geq 0$, and (ii) $\Delta(\beta, \delta) < 0$. For case (i), $\Delta(\gamma', \alpha) = \Delta(\beta, \delta) + \mathrm{n}$ for some natural number n. So let ϵ be any consequence such that $\Delta(\gamma', \epsilon) = \mathrm{n}$; then $\Delta(\epsilon, \alpha) = \Delta(\beta, \delta)$, which implies $\epsilon H \delta \sim \alpha H \beta$ under H3. The reasoning is more or less the same for case (ii). Finally, H1 implies that $\alpha\beta \sim_h \beta\alpha$ for all α, β, so the (b) parts of ACS3–ACS5 are each entailed by the respective (a) parts.

We can now make use of the following result:

Theorem 3 (Krantz et al. [9], 257) *Suppose that $\langle \mathbf{X} \times \mathbf{Y}, \succsim \rangle$ is an additive conjoint structure. Then there exists a pair of functions $u_x : \mathbf{X} \to \mathbb{R}$ and $u_y : \mathbf{Y} \to \mathbb{R}$ such that, for all $x, x' \in \mathbf{X}$ and $y, y' \in \mathbf{Y}$,*

$$xy \succsim x'y' \;\Leftrightarrow\; u_x(x) + u_y(y) \geq u_x(x') + u_y(y').$$

Furthermore, if u'_x and u'_y are two other functions with the same property, then there exist constants $\mathrm{a} > 0$, b_1 and b_2 such that $u'_x = \mathrm{a}u_x + \mathrm{b}_1$ and $u'_y = \mathrm{a}u_y + \mathrm{b}_2$.

In particular, let $u_x : \mathcal{C} \to \mathbb{R}$ and $u_y : \mathcal{C} \to \mathbb{R}$ by the two functions specified in the theorem. Since $\alpha H \beta \sim \beta H \alpha$ for all consequences α, β, it follows that

$$u_x(\alpha) + u_y(\beta) = u_x(\beta) + u_y(\alpha),$$

which is equivalent to

$$u_x(\alpha) - u_y(\alpha) = u_x(\beta) - u_y(\beta).$$

Thus $u_x = u_y + \mathrm{b}$ for some constant b. In light of this, H1 plus Theorem 3's uniqueness condition (i.e., with $\mathrm{a} = 1$, $\mathrm{b}_1 = 0$, $\mathrm{b}_2 = \mathrm{b}$) imply that there is a utility function u such that for all half-probability events H,

$$\alpha H \beta \succsim \gamma H \delta \;\Leftrightarrow\; u(\alpha) + u(\beta) \geq u(\gamma) + u(\delta) \;\Leftrightarrow\; u(\alpha) - u(\gamma) \geq u(\delta) - u(\beta).$$

So u satisfies U2. For U1, simply note:

$$\alpha \succsim \beta \;\Leftrightarrow\; \alpha H \alpha \succsim \beta H \beta \;\Leftrightarrow\; u(\alpha) + u(\alpha) \geq u(\beta) + u(\beta) \;\Leftrightarrow\; u(\alpha) \geq u(\beta).$$

The equal spacing property U3 follows from U2 plus H3. Finally, U4 follows straightforwardly from the uniqueness condition in Theorem 3. □

Lemma 3 *If C1–C8 hold, then there is a probability function p such that $\succsim$ maximizes expected utility relative to p.*

Proof Assume C1–C8. Let ω^0 be a consequence that's minimal in $\succsim$ over $\mathcal{C}$ (i.e., $\alpha \succsim \omega$ for all $\alpha \in \mathcal{C}$); and let ω^i be a consequence such that $\Delta(\omega^i, \omega^0) = i$. From C8 plus C5, there will exist a non-null state (call it $s_{\min}$) such that for every non-null state s there is a consequence (call it β_s) where $\beta_s \succsim \omega^1$ and $\omega^1\{s\}\omega^0 \sim \beta_s\{s_{\min}\}\omega^0$. Additionally, if there's any consequence $\gamma \neq \beta_s$ such that $\omega^1\{s\}\omega^0 \sim \gamma\{s_{\min}\}\omega^0$, then it follows from C7 that $\beta_s\{s\}\omega^0 \sim \gamma\{s\}\omega^0$, which by C6 implies $\gamma \sim \beta_s$ and so $\Delta(\beta_s, \omega^0) = \Delta(\gamma, \omega^0)$.

Given those facts, let p be the unique normalized function on $2^{\mathcal{S}}$ that satisfies the following conditions:

$$p(E) = \sum_{s \in E} p(s), \qquad p(s) = \begin{cases} 0 & \text{if } s \text{ is null,} \\ \Delta(\beta_s, \omega^0) \cdot p(s_{\min}) & \text{if } s \text{ is non-null.} \end{cases}$$

Observe that, since $\Delta(\beta_s, \omega^0)$ is always positive, $p(s_{\min})$ must be positive and p is non-negative, hence a probability function. Furthermore, let u be the unique utility function satisfying U1–U3 such that $u(\omega^i) = i$. Now p and u jointly induce a preference ordering (call it $\succsim^\star$) under the expected utility rule. We need to show the following:

R1. If $f \sim^\star g$, then $f \sim g$

R2. If $f \succ^\star g$ and there's no h such that $f \succ^\star h \succ^\star g$, then $f \succ g$

Since R1 and R2 together imply that $\succsim = \succsim^\star$, it will follow that $\succsim$ maximizes expected utility relative to p.

For R1, the goal will be to show that if $f \sim^\star g$, then $f \sim g$ follows by a sequence of indifferences justified by C5. To that end, it will be helpful to start by reconstructing the act space as a set of integer-valued vectors. Say that acts f and g are *essentially identical* (written $f \simeq g$) just in case $f(s) \sim g(s)$ for all non-null states s. Let $\mathbf{A}$ be the partition of $\mathcal{A}$ into $\simeq$-equivalence classes, so $\mathbf{f} \in \mathbf{A}$ just in case $\mathbf{f} = \{g \in \mathcal{A} \mid g \simeq f\}$ for some act f. Note that the definition of $\simeq$ doesn't depend on the choice of $\succsim$ or $\succsim^\star$, since s is null just in case $p(s) = 0$ and $\alpha \sim \beta \Leftrightarrow u(\alpha) = u(\beta) \Leftrightarrow \alpha \sim^\star \beta$ for all consequences α, β. For the same reason, $f \simeq g$ implies $f \sim g$ and $f \sim^\star g$. Furthermore, let $\widetilde{\mathcal{C}}$ be the set of $\sim$-equivalence classes in $\mathcal{C}$, and let $\mathcal{N} \subset \mathcal{S}$ be the set of null states. Since C3 implies there are half-probability events and C8 implies that $\succsim$ is non-trivial, it follows that $\mathcal{S} - \mathcal{N} = \{s_1, \ldots, s_{\mathrm{n}}\}$ for some $\mathrm{n} \geq 2$ and $|\widetilde{\mathcal{C}}| = \mathrm{m} + 1$ for some $\mathrm{m} \geq 1$. The upshot is that $\mathbf{A}$ is isomorphic to (and will therefore be identified with) the discrete n-dimensional box $\{0, 1, \ldots, \mathrm{m}\}^{\mathrm{n}}$. That is, every act f is represented by an n-tuple (or vector) $\mathbf{f}$ taking integer values between 0 and m inclusive, where the i^{th} coordinate expresses the utility of $f(s_i)$ under u. Two acts correspond to the same $\mathbf{f}$ just in case they're essentially identical, in which case they're indifferent under $\succsim$ and $\succsim^\star$.

Next we characterize the set of vector differences, $\mathbf{D} \subseteq \{-\mathrm{m}, \ldots, \mathrm{m}\}^{\mathrm{n}}$, such that $f \sim^\star g$ just in case $\mathbf{f} - \mathbf{g} \in \mathbf{D}$. First, where $\mathbb{Q}^{\mathrm{n}}$ is the vector space of functions from $\{1, \ldots, \mathrm{n}\}$ to $\mathbb{Q}$, define the linear map $eu : \mathbb{Q}^{\mathrm{n}} \to \mathbb{Q}$ such that for all $\mathbf{v} \in \mathbb{Q}^{\mathrm{n}}$,

$$eu(\mathbf{v}) \;=\; \sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot p(s_i).$$

The integer kernel of eu is then the lattice

$$\mathbf{K} \;=\; \ker(eu) \cap \mathbb{Z}^{\mathrm{n}} \;=\; \left\{ \mathbf{v} : \{1, \ldots, \mathrm{n}\} \to \mathbb{Z} \;\middle|\; \sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot p(s_i) = 0 \right\}.$$

Henceforth, assume for notational convenience that $s_{\mathrm{n}} = s_{\min}$, and let b_i be the positive integer ($\leq \mathrm{m}$) such that $p(s_i) = b_i \cdot p(s_{\min})$. Then, for each non-null $s_i \neq s_{\min}$, define the

basic exchange vector $\mathbf{b}^i_{\min}$ like so:

$$\mathbf{b}^i_{\min}(s) = \begin{cases} b_i & \text{if } s = s_{\min}, \\ -1 & \text{if } s = s_i, \\ 0 & \text{otherwise.} \end{cases}$$

Every basic exchange vector lies in $\mathbf{K}$, since

$$eu(\mathbf{b}^i_{\min}) \;=\; b_i \cdot p(s_{\min}) + (-1)p(s_i) \;=\; \frac{p(s_i)}{p(s_{\min})} \cdot p(s_{\min}) - p(s_i) \;=\; 0.$$

The same is true for every integer linear combination of basic exchange vectors, by the linearity of eu. Furthermore, every vector in $\mathbf{K}$ can be written as an integer linear combination of basic exchange vectors. Since $p(s_{\min}) > 0$,

$$\sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot p(s_i) \;=\; 0 \;\Leftrightarrow\; p(s_{\min}) \cdot \sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot b_i \;=\; 0 \;\Leftrightarrow\; \sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot b_i \;=\; 0.$$

And since $b_{\mathrm{n}} = 1$, isolating $\mathbf{v}(\mathrm{n}) \cdot b_{\mathrm{n}}$ from the final equality gets us:

$$\mathbf{v}(\mathrm{n}) \;=\; -\sum_{i=1}^{\mathrm{n}-1} \mathbf{v}(i) \cdot b_i.$$

Thus every kernel vector $\mathbf{v}$ takes the form:

$$\mathbf{v} \;=\; \sum_{i=1}^{\mathrm{n}-1} \big(-\mathbf{v}(i)\big) \cdot \mathbf{b}^i_{\min} \;=\; \left(\mathbf{v}(1), \ldots, \mathbf{v}(\mathrm{n}-1), -\sum_{i=1}^{\mathrm{n}-1} \mathbf{v}(i) \cdot b_i\right).$$

Now given $\mathbf{v}(i) \in \mathbb{Z}$ whenever $\mathbf{v} \in \mathbf{K}$, so $\mathbf{K}$ consists of the integer linear combinations of the $\mathbf{b}^i_{\min}$. Finally, $\mathbf{D}$ will be the intersection of $\mathbf{K}$ with those difference vectors that can in principle be realized from the elements of $\mathbf{A}$:

$$\mathbf{D} \;=\; \left\{ \sum_{i=1}^{\mathrm{n}-1} \mathrm{a}_i \cdot \mathbf{b}^i_{\min} \;\middle|\; \mathrm{a}_i \in \mathbb{Z} \right\} \cap \{-\mathrm{m}, \ldots, \mathrm{m}\}^{\mathrm{n}}.$$

We will also make use of the following observations. The first is just another way to express C5: two ordered pairs $(\mathbf{f}, \mathbf{g})$ and $(\mathbf{f}', \mathbf{g}')$ in $\mathbf{A}\times\mathbf{A}$ realize the same difference vector only if preferences between the acts they represent are matched. That is, for all $\mathbf{f}, \mathbf{g}, \mathbf{f}', \mathbf{g}' \in \mathbf{A}$,

C5. $\mathbf{f} - \mathbf{g} = \mathbf{f}' - \mathbf{g}'$ implies $f \succsim g \Leftrightarrow f' \succsim g'$

Next, for any $\mathbf{v}$, let $\mathbf{v}^{(+)}$ and $\mathbf{v}^{(-)}$ denote the pointwise positive and negative components of $\mathbf{v}$ respectively; so

$$\mathbf{v}^{(+)}(i) = \max\{\mathbf{v}(i), 0\}, \qquad \mathbf{v}^{(-)}(i) = \max\{-\mathbf{v}(i), 0\}.$$

If $\mathbf{f}, \mathbf{g} \in \mathbf{A}$ and $\mathbf{f} - \mathbf{g} = \mathbf{h}$, then $-\mathrm{m} \leq \mathbf{h}(i) \leq \mathrm{m}$ and so $\mathbf{h}^{(+)}$ and $\mathbf{h}^{(-)}$ also belong to $\mathbf{A}$. Moreover, $\mathbf{f} - \mathbf{g} = \mathbf{h}^{(+)} - \mathbf{h}^{(-)}$. Since $h^{(+)}$ and $h^{(-)}$ have the lowest utility values for each state consistent with the vector $\mathbf{f} - \mathbf{g}$, $(\mathbf{h}^{(+)}, \mathbf{h}^{(-)})$ is the *minimal realizer* of the difference vector realized by $(\mathbf{f}, \mathbf{g})$. By C5, the preference ordering between acts f and g matches the ordering of those acts represented by the minimal realizer of $\mathbf{f} - \mathbf{g}$. That is, for all $\mathbf{f}, \mathbf{g} \in \mathbf{A}$,

C5a. $\mathbf{f} - \mathbf{g} = \mathbf{h}$ implies $f \succsim g \Leftrightarrow h^{(+)} \succsim h^{(-)}$

Recall also that for every non-null state s there's a consequence β_s where $\beta_s \succsim \omega^1$ and $\omega^1\{s\}\omega^0 \sim \beta_s\{s_{\min}\}\omega^0$. Since $u(\beta_{s_i}) = b_i$,

$$\beta_{s_i}\{s_{\min}\}\omega^0 \in \mathbf{b}^i_{\min}{}^{(+)}, \qquad \omega^1\{s_i\}\omega^0 \in \mathbf{b}^i_{\min}{}^{(-)}.$$

From C5$_a$,

$$\mathbf{f} - \mathbf{g} = \mathbf{b}^i_{\min} \;\Rightarrow\; \big(f \sim g \;\Leftrightarrow\; \beta_s\{s_{\min}\}\omega^0 \sim \omega^1\{s\}\omega^0\big),$$

in which case the known indifference $\omega^1\{s\}\omega^0 \sim \beta_s\{s_{\min}\}\omega^0$ implies $f \sim g$. The same holds if $\mathbf{f} - \mathbf{g} = -\mathbf{b}^i_{\min}$. Rearranging terms gets us, for all $\mathbf{f} \in \mathbf{A}$,

C5$_b$. $\mathbf{f} \pm \mathbf{b}^i_{\min} = \mathbf{g}$ implies $f \sim g$ whenever $\mathbf{g} \in \mathbf{A}$

Finally, suppose that $f \gtrsim g$ and that there are f', g' such that $\mathbf{f}-\mathbf{g} = \mathbf{f}'-\mathbf{g}'$ and $\mathbf{f}' \pm \mathbf{b}^i_{\min} = \mathbf{h}$ for some $\mathbf{h} \in \mathbf{A}$. By C5, $f' \gtrsim g'$; and by C5$_b$, $f' \sim h$; hence $h \gtrsim g'$. Now observe that

$$\mathbf{h} - \mathbf{g}' = \begin{cases} \big(\mathbf{f} + \mathbf{b}^i_{\min}{}^{(+)}\big) - \big(\mathbf{g} + \mathbf{b}^i_{\min}{}^{(-)}\big) & \text{if } \mathbf{f}' + \mathbf{b}^i_{\min} = \mathbf{h}, \\ \big(\mathbf{f} + \mathbf{b}^i_{\min}{}^{(-)}\big) - \big(\mathbf{g} + \mathbf{b}^i_{\min}{}^{(+)}\big) & \text{if } \mathbf{f}' - \mathbf{b}^i_{\min} = \mathbf{h}. \end{cases}$$

Consequently, for all $\mathbf{f}, \mathbf{g} \in \mathbf{A}$,

C5$_c$. $\mathbf{f} + \mathbf{b}^i_{\min}{}^{(+)} = \mathbf{h}$ and $\mathbf{g} + \mathbf{b}^i_{\min}{}^{(-)} = \mathbf{h}'$, or $\mathbf{f} + \mathbf{b}^i_{\min}{}^{(-)} = \mathbf{h}$ and $\mathbf{g} + \mathbf{b}^i_{\min}{}^{(+)} = \mathbf{h}'$, implies $f \gtrsim g \Leftrightarrow h \gtrsim h'$ whenever $\mathbf{h}, \mathbf{h}' \in \mathbf{A}$

Putting all that together: if $f \sim^\star g$, then $\mathbf{f} - \mathbf{g} = \mathbf{h}$ for some $\mathbf{h} \in \mathbf{D}$, and so $f \sim g \Leftrightarrow h^{(+)} \sim h^{(-)}$. To establish R1 it therefore suffices to show that $h^{(+)} \sim h^{(-)}$. Keeping in mind the earlier stipulation that $s_n = s_{\min}$, recall that $\mathbf{b}^i_{\min}(i) = -1$ and $\mathbf{b}^i_{\min}(j) = 0$ whenever $i \neq j \in \{1, \ldots, \mathrm{n}-1\}$. As such, let $\mathbf{h}_{[n-1]}$ be the restriction of $\mathbf{h}$ to the first $n-1$ coordinates, so

$$\mathbf{h} = (x_1, \ldots, x_{n-1}, x_n) \;\Rightarrow\; \mathbf{h}_{[n-1]} = (x_1, \ldots, x_{n-1}).$$

Then $\mathbf{h}_{[n-1]}(i)$ expresses the number of times ($\leq m$) the corresponding basic exchange vectors $\mathbf{b}^i_{\min}$ must be added (where $\mathbf{h}(i) > 0$) or subtracted (where $\mathbf{h}(i) < 0$) from $\mathbf{h}$ to obtain the zero vector (designated $\mathbf{0}$). For example, if $\mathbf{h} = (-2, 3, 4)$, then $\mathbf{h}_{[n-1]} = (-2, 3)$; and if $\mathbf{h} \in \mathbf{D}$, it follows that

$$\mathbf{h} - 2\mathbf{b}^1_{\min} + 3\mathbf{b}^2_{\min} = \mathbf{0}.$$

That is, we obtain $\mathbf{0}$ by taking $\mathbf{b}^1_{\min}$ from $\mathbf{h}$ twice and adding $\mathbf{b}^2_{\min}$ three times. We can then show that $h^{(+)} \sim h^{(-)} \Leftrightarrow f \sim f$ by chaining together indifferences justified by the foregoing observations. This amounts to a finite 'walk' through the space $\mathbf{A}\times\mathbf{A}$, from $(\mathbf{h}^{(+)}, \mathbf{h}^{(-)})$ to $(\mathbf{0}, \mathbf{0})$, with steps corresponding to the addition or subtraction of a basic exchange vector to/ from $\mathbf{h}$.

Given the earlier decomposition of $\mathbf{h} \in \Delta$,

$$\mathbf{h} = \sum_{i=1}^{n-1} a_i \cdot \mathbf{b}^i_{\min}, \quad a_i \in \mathbb{Z},$$

set the *positive* and *negative exchange multisets* for $\mathbf{h}$ like so:[20]

$$\mathrm{Pos}(\mathbf{h}) \;=\; \big\{i^{\langle a_i \rangle} \,|\, a_i > 0\big\}, \qquad \mathrm{Neg}(\mathbf{h}) \;=\; \big\{i^{\langle |a_i| \rangle} \,|\, a_i < 0\big\}.$$

For any pair $(\mathbf{f}, \mathbf{g})$, call $\mathbf{f}$ the *top vector* and $\mathbf{g}$ the *bottom vector*. We take a *+i-step* from $(\mathbf{f}, \mathbf{g})$ when we add $\mathbf{b}^i_{\min}{}^{(+)}$ to the top vector and add $\mathbf{b}^i_{\min}{}^{(-)}$ to the bottom vector, arriving at $(\mathbf{f} + \mathbf{b}^i_{\min}{}^{(+)}, \mathbf{g} + \mathbf{b}^i_{\min}{}^{(-)})$. Similarly, we take a *−i-step* from $(\mathbf{f}, \mathbf{g})$ when we add $\mathbf{b}^i_{\min}{}^{(-)}$ to the top and $\mathbf{b}^i_{\min}{}^{(+)}$ to the bottom vectors. A +i-step (or −i-step) is *available* just in case there is an i in $\mathrm{Pos}(\mathbf{h})$ (or $\mathrm{Neg}(\mathbf{h})$, respectively), and with each such step we eliminate an i from that multiset.[21] A ±i-step *stays in bounds* whenever the coordinates of the resulting vectors

[20] Here, $i^{\langle n \rangle}$ denotes a multiset (also known as a *bag*) comprising n copies of i. For example, $7^{\langle 3 \rangle} = \{7, 7, 7\}$ and $2^{\langle 3 \rangle}3^{\langle 2 \rangle} = \{2, 2, 2, 3, 3\}$; and if $\mathbf{h}_{[n-1]} = (4, -3, 1)$, then $\mathrm{Pos}(\mathbf{h}) = \{1, 1, 1, 3\}$ and $\mathrm{Neg}(\mathbf{h}) = \{2, 2, 2\}$.

[21] For example, if $\mathbf{h}_{[n-1]} = (-2, 3)$, then at the start of the walk a total of two −1-steps and three +2-steps will be available; and if a single +2-step is taken, then two −1-steps and two +2-steps will remain available.

stay within $\{0, \ldots, m\}$, and it's *valid* just in case it's available and stays in bounds. Finally, we take a *reduction step* (or r-step) by shifting from $(\mathbf{f}, \mathbf{g})$ to $\big((\mathbf{f}-\mathbf{g})^{(+)}, (\mathbf{f}-\mathbf{g})^{(-)}\big)$—i.e., reducing to the minimal realizer of the corresponding difference vector.

By C5$_a$ and C5$_c$, if we start at any point $(\mathbf{f}, \mathbf{g})$ in $\mathbf{A}\times\mathbf{A}$ and then take any valid $\pm i$-step followed immediately by an r-step, we will land on a pair $(\mathbf{f}', \mathbf{g}')$ such that

$$\mathbf{f}' - \mathbf{g}' = (\mathbf{f} - \mathbf{g}) \pm \mathbf{b}^i_{\min} \in \mathbf{A},$$

and therefore $f \sim g \Leftrightarrow f' \sim g'$. Consequently, starting at $(\mathbf{h}^{(+)}, \mathbf{h}^{(-)})$ for any $\mathbf{h} \in \mathbf{D}$, we apply the following procedure:

S1. If there are any available $\pm i$-steps in $\mathrm{Pos}(\mathbf{h})$ or $\mathrm{Neg}(\mathbf{h})$, then go to S2; otherwise, terminate procedure.

S2. If there are no valid $+i$-steps in $\mathrm{Pos}(\mathbf{h})$, then move to S3; otherwise, take any valid $+i$-step followed by an r-step, then go to S1.

S3. If there are no valid $-i$-steps $\mathrm{Neg}(\mathbf{h})$, then move to S1; otherwise, take any valid $-i$-step followed by an r-step, then go to S1.

By the points above, if all available $\pm i$-steps are taken then the procedure will terminate at $(\mathbf{0}, \mathbf{0})$, with every $\pm i$-step and r-step along the way landing on some $(\mathbf{f}, \mathbf{g}) \in \mathbf{A}\times\mathbf{A}$ such that $h^{(+)} \sim h^{(-)} \Leftrightarrow f \sim g$. So to complete the proof of R1 we just need to show that all available $\pm i$-steps will indeed be taken—that is,

(i) following any r-step, if there are no valid $+i$-steps but there are some available $-i$-steps, then at least one $-i$-step stays in bounds; and

(ii) following any r-step, if there are no valid $-i$-steps but there are some available $+i$-steps, then at least one $+i$-step stays in bounds.

We focus on showing (i), since the case for (ii) is symmetric. Consider to start with the first $n-1$ coordinates. There are $\mathbf{h}^{(-)}(i)$ $-i$-steps available at the beginning of the walk, and $\mathbf{h}^{(+)}(i) = 0$ whenever $\mathbf{h}^{(-)}(i) > 0$. Furthermore, a $-i$-step adds 1 to the i^{th} coordinate of the top vector, otherwise leaving the first $n-1$ coordinates of both the top and bottom vectors untouched. So the net effect of a $-i$-step followed by an r-step on those coordinates is a reduction by 1 of the i^{th} coordinate of the bottom vector, leaving the i^{th} coordinate of the top vector at zero; and furthermore, taking $\mathbf{h}^{(-)}(i)$ $-i$-steps results in both top and bottom vectors taking value 0 at coordinate i. (The same is true for $+i$-steps, *mutatis mutandis*.) As such, an available $-i$-step never goes out of bounds with respect to the first $n-1$ coordinates. Restricting our attention to the n^{th} coordinate, then, observe that a $+i$-step always adds $1 \leq b_i \leq m$ to coordinate n at the top vector, and a negative j-step always adds $1 \leq b_j \leq m$ to coordinate n at the bottom vector. There are two possible circumstances where there are no valid $+i$-steps following an r-step:

1. There are no available $+i$-steps, and
2. There is an available $+i$-step but it goes out of bounds.

If it's the latter, then the top vector must be positive at n, else a $+i$-step would not go out of bounds. If it's the former, and there are some negative j-steps available, then by the points above the top vector must be zero at the first $n-1$ coordinates and the bottom vector must be positive for at least one of those coordinates. Since every vector in the kernel of eu besides $\mathbf{0}$ must have non-zero positive and negative components, and since every step in the walk lands on a pair of vectors that realizes a kernel vector, it follows again that the top vector must be positive at n. So in both circumstances, the top vector is positive at n, in which case the bottom vector must be zero at n, and so taking any available $-i$-step will stay in bounds.

For R2, recall that for all non-null states s_i, $\mathbf{f}(i) = u\big(f(s_i)\big)$ and $p(s_i) = b_i \cdot p(s_{\min})$ for some integer $1 \leq b_i \leq m$. Given this,

$$eu(\mathbf{f}) \;=\; p(s_{\min}) \cdot \sum_{i=1}^{\mathrm{n}} b_i \cdot \mathbf{f}(i) \;=\; \sum_{s \in \mathcal{S}} p(s) \cdot u\big(f(s)\big).$$

Since eu represents $\succsim^\star$, we know that

$$f \succ^\star g \;\Leftrightarrow\; eu(\mathbf{f}) - eu(\mathbf{g}) \geq p(s_{\min}).$$

Let $\mathbf{e}_{\min}$ be the vector assigning $+1$ for $s_{\min}$ and 0 otherwise. Thus $\mathbf{f} - \mathbf{g} = \mathbf{e}_{\min}$ when $eu(\mathbf{f}) - eu(\mathbf{g}) = p(s_{\min})$, in which case $f \succ^\star g$ and (moreover) there can be no h such that $f \succ^\star h \succ^\star g$. Furthermore, let g be any act that's not maximal in $\succsim^\star$. Then there is an act g' such that $g \sim^\star g'$ and $\mathbf{g}' + \mathbf{e}_{\min} = \mathbf{f}$ for some act f:

1. If $u\big(g(s_{\min})\big) < \mathrm{m}$, then $\mathbf{g}' = \mathbf{g}$ and $\mathbf{g}' + \mathbf{e}_{\min} \in \mathbf{A}$.

2. If $u\big(g(s_{\min})\big) = \mathrm{m}$, then $\mathbf{g} + \mathbf{e}_{\min} \notin \mathbf{A}$. But since g is non-maximal, there must be some non-null state $s_i \neq s_{\min}$ such that $u\big(g(s_i)\big) < \mathrm{m}$. So $\mathbf{g}' = \mathbf{g} - \mathbf{b}^i_{\min} \in \mathbf{A}$ and $\mathbf{g}' + \mathbf{e}_{\min} \in \mathbf{A}$.

As such, suppose that $f \succ^\star g$ and there's no h such that $f \succ^\star h \succ^\star g$. Then there will be acts f' and g' such that $f \sim^\star f'$, $g \sim^\star g'$ and $\mathbf{f} - \mathbf{g}' = \mathbf{e}_{\min} - \mathbf{0}$. By C5$_a$ and C6, $\mathbf{f} - \mathbf{g}' = \mathbf{e}_{\min} - \mathbf{0}$ implies $f \succ g'$; by R1, $f \sim f'$ and $g \sim g'$; therefore $f \succ g$. □

From here the proof of Theorem 2 is straightforward. From Lemma 3 we get PGR1: $\succsim$ maximizes expected utility relative to some probability function p. Given the conclusions established in Section 4, for PGR3 we just need to show that $\succsim$ maximizes expected utility relative to c just in case $c = \mathrm{a} \cdot p$ for some non-zero constant a. The right-to-left direction is obvious. For the left-to-right, we note that if $c \neq \mathrm{a} \cdot p$ for any non-zero a, then c is either (i) constant ($\mathrm{a} = 0$); or, if it's non-constant, then (ii) it's non-additive or (iii) inconsistent with the basic exchange rates between states. Since all three cases can be ruled out, p must be the unique probability function p such that $\succsim$ maximizes expected utility relative to p. Finally, PGR2 follows immediately from C2 given PGR1.

References

[1] Narens, L.: On the scales of measurement. Journal of Mathematical Psychology **24**(3), 249–275 (1981)

[2] Joyce, J.: The Value of Truth: A Reply to Howson. Analysis **75**(3), 413–424 (2015)

[3] Finetti, B.: Sul Significato Soggettivo Della Probabilita. Fundamenta Mathematicae **17**(1), 298–329 (1931)

[4] Koopman, B.O.: The Axioms and Algebra of Intuitive Probability. Annals of Mathematics **41**(2), 269–292 (1940)

[5] Koopman, B.O.: The Bases of Probability. Bulletin of the American Mathematical Society **46**(10), 763–774 (1940)

[6] Savage, L.J.: The Foundations of Statistics. Dover, New York (1954)

[7] Villegas, C.: On Qualitative Probability σ-Algebras. The Annals of Mathematical Statistics **35**(4), 1787–1796 (1964)

[8] Luce, D.R.: On the numerical representation of qualitative conditional probability. Annals of Mathematical Statistics **39**, 481–491 (1968)

[9] Krantz, D.H., Luce, R.D., Suppes, P., Tversky, A.: Foundations of Measurement, Vol. I: Additive and Polynomial Representations. Academic Press, New York (1971)

[10] Fine, T.L.: Theories of Probability: An Examination of Foundations. Academic Press, New York (1973)

[11] Suppes, P., Zanotti, M.: Necessary and sufficient conditions for existence of a unique measure strictly agreeing with a qualitative probability ordering. Journal of Philosophical Logic **5**(3), 431–438 (1976)

[12] Suppes, P., Zanotti, M.: Necessary and Sufficient Qualitative Axioms for Conditional Probability. Z. Wahrschelnllchkeltstheorle verw. Gebiete **60**, 163–169 (1982)

[13] Suppes, P.: Qualitative theory of subjective probability. In: Wright, G., Ayton, P. (eds.) Subjective Probability, pp. 17–38. Wiley, New York (1994)

[14] Zynda, L.: Representation Theorems and Realism About Degrees of Belief. Philosophy of Science **67**(1), 45–69 (2000)

[15] Suppes, P.: Qualitative Axioms of Uncertainty as a Foundation for Probability and Decision-Making. Minds and Machines **26**, 185–202 (2016)

[16] DiBella, N.: Qualitative Probability and Infinitesimal Probability (2018)

[17] Stefánsson, H.O.: On the Ratio Challenge for Comparativism. Australasian Journal of Philosophy **96**(2), 380–390 (2018)

[18] Konek, J.: Comparative Probabilities. In: The Open Handbook of Formal Epistemology, pp. 267–348. PhilPapers Foundation, n.p. (2019)

[19] Elliott, E.: The Measurement of Subjective Probability. Cambridge University Press, New York (2024)

[20] Neth, S.: Better Foundations for Subjective Probability. Australasian Journal of Philosophy **103**(1) (2025)

[21] Christensen, D.: Preference-based arguments for probabilism. Philosophy of Science **68**(3), 356–376 (2001)

[22] Eriksson, L., Hájek, A.: What are degrees of belief? Studia Logica **86**(2), 183–213 (2007)

[23] Hájek, A.: Arguments for—or against—Probabilism? British Journal for the Philosophy of Science **59**(4), 793–819 (2008)

[24] Meacham, C.J.G., Weisberg, J.: Representation Theorems and the Foundations of Decision Theory. Australasian Journal of Philosophy **89**(4), 641–663 (2011)

[25] Titelbaum, M.: Fundamentals of Bayesian Epistemology 2: Arguments, Challenges, Alternatives. Oxford University Press, New York (2022)

[26] Savage, L.J.: Implications of Personal Probability for Induction. The Journal of Philosophy **64**(19), 593–607 (1967)

[27] Savage, L.J.: Elicitation of Personal Probabilities and Expectations. Journal of the American Statistical Association **66**(336), 783–801 (1971)

[28] Joyce, J.: The Foundations of Causal Decision Theory. Cambridge University Press, New York (1999)

[29] Jeffrey, R.: Frameworks for Preference. In: Balch, M., McFadden, D., Wu, S. (eds.) Essays on Economic Behavior Under Uncertainty. North-Holland Publishing Company, Amsterdam (1974)

[30] Ramsey, F.P.: Truth and probability. In: Braithwaite, R.B. (ed.) The Foundations of Mathematics and Other Logical Essays, pp. 156–198. Routledge, New York (1931)

[31] Suppes, P.: The Measurement of Belief. Journal of the Royal Statistical Society. Series B (Methodological) **36**(2), 160–191 (1974)

[32] Shafer, G.: Savage revisited. Statistical Science **1**(4), 463–501 (1986)

[33] Gilboa, I.: Theory of Decision Under Uncertainty. Cambridge University Press, New York (2013)

[34] Abdellaoui, M., Wakker, P.: Savage for dummies and experts. Journal of Economic Theory **186**, 1–20 (2020)

[35] Davidson, D., Suppes, P.: A finitistic axiomatization of subjective probability and utility. Econometrica **24**(3), 264–275 (1956)

[36] Pfanzagl, J.: Theory of Measurement. Wiley, New York (1968)

[37] Luce, R.D.: Dimensionally Invariant Numerical Laws Correspond to Meaningful Qualitative Relations. Philosophy of Science **45**(1), 1–16 (1978)

[38] Narens, L.: Abstract Measurement Theory. MIT Press, Cambridge, MA (1985)

[39] Roberts, F.: Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences. Cambridge University Press, New York (1985)

[40] Falmagne, J.C., Narens, L.: Scales and Meaningfulness of Quantitative Laws. Synthese **55**, 287–325 (1983)

[41] Luce, R.D., Krantz, D.H., Suppes, P., Tversky, A.: Foundations of Measurement, Vol. III: Representation, Axiomatization, and Invariance. Dover, New York (1990)

[42] Mundy, B.: On the general theory of meaningful representation. Synthese **67**, 391–437 (1986)

[43] Hawthorne, J.: A Logic of Comparative Support: Qualitative Conditional Probability Relations Representable by Popper Functions. In: Hájek, A., Hitchcock, C. (eds.) Oxford Handbook of Probabilities and Philosophy. Oxford University Press, Oxford (2016)

[44] Domotor, Z.: Qualitative information and entropy structures. In: Hintikka, J., Suppes, P. (eds.) Information and Inference, pp. 148–194. Reidel, New York (1970)

[45] Kaplan, M., Fine, T.L.: Joint orders in comparative probability. The Annals of Probability **5**, 161–179 (1977)

[46] Luce, R.D., Narens, L.: Qualitative independence in probability theory. Theory and Decision **9**, 225–239 (1978)

[47] Joyce, J.: A Defense of Imprecise Credences in Inference and Decision Making. Philosophical Perspectives **24**(1), 281–323 (2010)