Realism and Representations in Decision Theory

Abstract

Under what conditions does an agent have credences that can be represented probabilistically? Is it just a matter of having the right preferences, or something deeper? Building on an insight due to Zynda, this paper develops an account of when two decision-theoretic representations count as different ways of saying the same thing, even if they represent credences very differently. The account is applied to representations in non-additive variants of expected utility theory, rank-dependent utility theory, and state-dependent utility theory—all of which determine the same preferences, but not all of which are psychologically equivalent.

1. Introduction

At the heart of (subjective) expected utility theory sits a biconditional, linking preferences on one side to probabilities and utilities on the other. In words, it asserts that an agent prefers one act to another if and only if having that preference maximizes expected utility relative to some probability and utility function.

Within a finite version of Savage's (1954) framework, we can express the theory as follows. Let $\mathcal{S}$ be a finite set of *states*, sets of which are called *events*, and let $\mathcal{C}$ be a finite set of *consequences*. An *act* is any function from states to consequences. For any act f and event E on which f is constant, f_E denotes the consequence assigned by f throughout E; and $\mathcal{E}_f$ denotes a partition of $\mathcal{S}$ on which f is constant, so $E \in \mathcal{E}_f$ implies $f(s) = f_E$ for all $s \in E$. The basic rule of expected utility theory then requires that for all acts f, g and any choice of corresponding partitions $\mathcal{E}_f, \mathcal{E}_g$,

$$f \succsim g \Leftrightarrow \sum_{E \in \mathcal{E}_f} p(E)u(f_E) \geq \sum_{E \in \mathcal{E}_g} p(E)u(g_E),$$

where $\succsim$ is a weak preference relation over acts, u is a real-valued utility function on $\mathcal{C}$, and p is a probability function on the algebra of events induced by $\mathcal{S}$.

This formalization is consistent with several interpretations. We might read the biconditional *descriptively* (as characterizing what ordinary agents are usually like), *normatively* (as characterizing how rational agents ought to be), or *constitutively* (as characterizing what it is to be an agent). We might also interpret $\succsim$ *behaviorally* (as representing choice dispositions) or *mentally* (as representing some comparative desire-like attitude). But I flag these distinctions only to note that they shouldn't matter here. I use descriptive language throughout for ease of expression; it should be easy enough to translate if you favor some other interpretation.

My topic involves a different interpretive question, concerning the explanatory relationship between the preferences on the left and the probabilities on the right. In broad strokes, *constructivists* take preferences to be fundamental, with probabilities being no more than an interpretive 'construction' for representing patterns in a coherent preference ordering. Hence constructivists see an explanatory arrow pointing left-to-right: it's the structure of $\succsim$, together with our representational conventions, that explains why p takes the shape it does. *Realists*, by contrast, interpret p as representing an agent's credences, where these (i) have some psychological reality independent of, or at least not reducible to, preferences, and (ii) help causally determine preferences. So realists see an explanatory arrow pointing right-to-left: the agent's credences, represented by p, help explain the structure of $\succsim$.

As neat as that division might seem, there's more than one way to interpret a probability function as representing something 'psychologically real'. As Zynda (2000) pointed out, whenever an agent's preferences can be seen as maximizing expected utility relative to credences represented by some probability function, those same preferences can also be seen as maximizing some other aggregate value relative to credences represented non-probabilistically. The question for realists is whether these alternative representations express genuinely different psychological models, or whether they're just notational variants of the usual probabilistic representation—and without an answer, it's unclear what the probabilistic representation of credences even amounts to.

In this paper, I expand on Zynda's example and propose an account of when two decision-theoretic representations amount to different ways of saying the same thing. The goal is to sharpen what realism is—and is not—committed to. The bulk of the discussion will be framed around how realists should respond to a range of alternative representation schemes, described in §2. In §3, I sketch three species of realism—*numerical*, *scalar* and *structural*—with structural realism developed further in §4. Finally, in §5, I clarify the core commitments of structural realism by contrasting it with Zynda's *weak realism*.

2. Biconditionals galore!

It will pay to be systematic, so start with terminology. A *utility function* is any real-valued function on $\mathcal{C}$, and a *credence function* is any real-valued function on the event algebra $2^{\mathcal{S}}$. Thus a credence function may take values outside the usual $[0, 1]$ range. A *probability function*, by contrast, is *strictly* defined as any credence function c satisfying the usual Kolmogorov properties:

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

Next, a *valuation* is to be understood as any real-valued function on $\mathcal{C}^{\mathcal{S}}$. An *expected utility valuation* (eu) is defined relative to a choice of (not necessarily probabilistic) credence function c and utility function u such that

$$eu^{(c,u)}(f) = \sum_{E \in \mathcal{E}_f} c(E)u(f_E),$$

independent of the choice of partition $\mathcal{E}_f$. Note that while c need not be a probability function, this definition is coherent only if c and u satisfy certain implicit constraints. In particular, partition independence implies that c must be additive. The non-*eu* valuations below will likewise impose constraints on their parameters. This will come up again later; for now, just keep in mind that a valuation need not be defined for every combination of credence and utility functions.

Finally, suppose Bayesia is a hypothetical agent whose preferences are represented by $\succsim$. We assume there exists a unique probability function p and some utility function u such that $\succsim$ *maximizes eu* relative to p and u, in the sense that

$$f \succsim g \Leftrightarrow eu^{(p,u)}(f) \geq eu^{(p,u)}(g). \tag{MEU}$$

The existence of this unique probability function amounts to a non-trivial assumption about Bayesia's preferences.[1] But it also isn't essential to the main points of discussion. See it as a framing device—the idea is that *even if* Bayesia's preferences make p 'unique' in this sense, p is not unique in the more general sense of being the *only* credence function that can be combined with *some* utility function under *some* valuation to generate those same preferences. I will give five examples of this, the first four of which are represented in Figure 1.

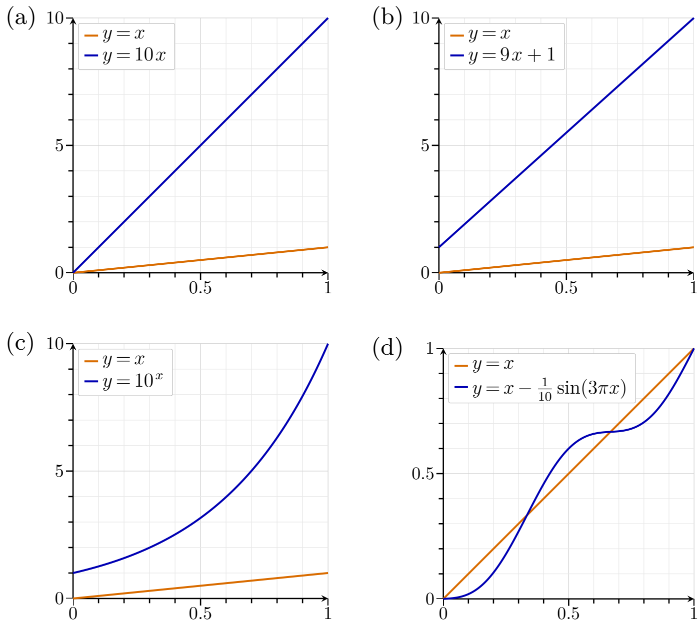

Figure 1: Four transformations of p, shown against the identity line $y = x$. The horizontal axis gives possible values of p, while the blue curve gives the corresponding transformed values. (a) $c_{\text{add}} = 10p$, a positive scalar transformation; (b) $c_{\text{aff}} = 9p+1$, a positive affine transformation; (c) $c_{\text{mul}} = 10^p$, a positive exponential transformation; (d) $c_{\text{rdu}} = p - \frac{1}{10}\sin(3\pi p)$, a representative order-preserving transformation.

[1] Savage's representation theorem supplies conditions under which preferences maximize expected utility relative to a unique probability function, though he requires infinite states. There are similar results for finite $\mathcal{S}$—see (Gul 1992) and (Abdellaoui and Wakker 2020).

2.1 Scalar transformations

The first example is the easiest. Suppose we take the 'unique' probability function p and apply some scalar transformation $x \mapsto \alpha x$ pointwise, with $\alpha > 0$—say

$$c_{\text{add}}(E) = 10p(E).$$

While c_{add} is not a probability function, it retains the same basic structure just stretched out by a factor of 10. Thus it satisfies:

1. $c_{\text{add}}(\mathcal{S}) = 10$ (renormalization)
2. $c_{\text{add}}(E) \geq 0$ (non-negativity)
3. $E \cap F = \varnothing$ implies $c_{\text{add}}(E \cup F) = c_{\text{add}}(E) + c_{\text{add}}(F)$ (additivity)

More importantly, observe that

$$eu^{(c_{\text{add}},u)}(f) = \sum_{E \in \mathcal{E}_f} 10p(E)u(f_E) = 10\, eu^{(p,u)}(f).$$

Since multiplying expected utilities by a positive constant preserves their ordering, it follows that (MEU) holds just in case

$$f \succsim g \iff eu^{(c_{\text{add}},u)}(f) \geq eu^{(c_{\text{add}},u)}(g). \quad \text{(ADD)}$$

In other words, Bayesia's preferences maximize eu relative to p and u if and only if they maximize eu relative to c_{add} and u.

2.2 Affine transformations

The second example is Zynda's. This time, we take p and apply an affine transformation $x \mapsto \alpha x + \beta$ pointwise, with $\alpha > 0$. Nothing crucial turns on the particular choice of α and β, but suppose

$$c_{\text{aff}}(E) = 9p(E) + 1.$$

Think of c_{aff} as p stretched out (by a factor of 9) and shifted upwards (by one unit). The result is that c_{aff} satisfies:

1. $c_{\text{aff}}(\mathcal{S}) = 10$ (renormalization)
2. $c_{\text{aff}}(E) \geq 1$ (unit lower-bounded)
3. $E \cap F = \varnothing$ implies $c_{\text{aff}}(E \cup F) = c_{\text{aff}}(E) + c_{\text{aff}}(F) - 1$ (affine-additivity)

Since c_{aff} is non-additive, $\succsim$ cannot maximize eu relative to c_{aff}. But that just means we need to tinker with the form of the valuation a little. Let an *affine-expected utility valuation* (*aff*) be defined relative to a choice of c and u such that for any $\mathcal{E}_f$,

$$aff^{(c,u)}(f) = \sum_{E \in \mathcal{E}_f} (c(E)u(f_E) - u(f_E)).$$

It follows that (MEU) holds just in case

$$f \succsim g \iff aff^{(c_{\text{aff}},u)}(f) \geq aff^{(c_{\text{aff}},u)}(g). \quad \text{(AFF)}$$

Proof:

$$\begin{aligned} \mathit{aff}^{(c_{\text{aff}},u)}(f) &= \sum_{E \in \mathcal{E}_f} ((9p(E)+1)u(f_E) - u(f_E)) \\ &= \sum_{E \in \mathcal{E}_f} 9p(E)u(f_E) \\ &= 9\,\mathit{eu}^{(p,u)}(f). \end{aligned}$$

So $\mathit{aff}^{(c_{\text{aff}},u)}$ is just $\mathit{eu}^{(p,u)}$ scaled by a positive constant; hence Bayesia's preferences maximize *eu* relative to p and u if and only if they maximize *aff* relative to c_{aff} and u.

2.3 Exponential transformations

Both c_{add} and c_{aff} are linearly related to p, but do not be misled into thinking that this is essential. This next example makes that clear, since it involves a credence function that's non-linearly related to p. That is, let c_{mul} be obtained from p by applying a positive exponential transformation $x \mapsto \alpha^x$ pointwise, with $\alpha > 1$. Let's say $\alpha = 10$, so

$$c_{\text{mul}}(E) = 10^{p(E)}.$$

Whereas an affine transformation stretches and shifts p's values, this transformation stretches, shifts, and then also bends those values—with the result being that c_{mul} satisfies:

1. $c_{\text{mul}}(\mathcal{S}) = 10$ (renormalization)
2. $c_{\text{mul}}(E) \geq 1$ (unit lower-bounded)
3. $E \cap F = \varnothing$ implies $c_{\text{mul}}(E \cup F) = c_{\text{mul}}(E) \cdot c_{\text{mul}}(F)$ (multiplicativity)

Since c_{mul} isn't additive, Bayesia's preferences cannot maximize *eu* relative to c_{mul}; and since it's not affine-additive, they cannot maximize *aff* relative to c_{mul} either. Again, though, this just means we need to construct a new valuation. Let the *multiplicative-expected utility valuation* (*mul*) be defined relative to a choice of c and u such that for any $\mathcal{E}_f$,

$$\mathit{mul}^{(c,u)}(f) = \prod_{E \in \mathcal{E}_f} c(E)^{u(f_E)}.$$

Now (MEU) holds just in case

$$f \succsim g \Leftrightarrow \mathit{mul}^{(c_{\text{mul}},u)}(f) \geq \mathit{mul}^{(c_{\text{mul}},u)}(g). \quad \text{(MUL)}$$

To see this, note first the following identities: for $n > 1$ and $y > 0$,

$$\exp_n(x \log_n y) = y^x$$

and

$$\exp_n\Big(\sum_i x_i\Big) = \prod_i \exp_n(x_i).$$

As such, given $p(E) = \log_{10} c_{\text{mul}}(E)$ we derive

$$\begin{aligned} mul^{(c_{\text{mul}},u)}(f) &= \prod_{E \in \mathcal{E}_f} c_{\text{mul}}(E)^{u(f_E)} \\ &= \prod_{E \in \mathcal{E}_f} \exp_{10}(u(f_E) \log_{10} c_{\text{mul}}(E)) \\ &= \exp_{10}\Big(\sum_{E \in \mathcal{E}_f} u(f_E) \log_{10} c_{\text{mul}}(E) \Big) \\ &= \exp_{10}(eu^{(p,u)}(f)). \end{aligned}$$

Finally, since the exponentiation is strictly increasing, $mul^{(c_{\text{mul}},u)}$ and $eu^{(p,u)}$ induce the same ordering over acts. So Bayesia's preferences maximize eu relative to p and u if and only if they maximize mul relative to c_{mul} and u.

2.4 Order-preserving transformations

The next example employs the same kind of valuation found in Choquet expected utility theory (Schmeidler 1989), with decision weights that decompose into a credence function and a risk function, as in rank-dependent utility theories (e.g., Quiggin 1982; Buchak 2014). A fair bit of setup is needed for later precision, but the gist is very straightforward: simply transform p into some other credence function, then just select the risk function that transforms it back into p.

Let c_{rdu} be any non-trivial order-preserving transformation of p with fixed endpoints, in the sense that

1. $c_{\text{rdu}} \neq p$ (non-triviality)
2. $p(E) = 0 \Rightarrow c_{\text{rdu}}(E) = 0$, and $p(E) = 1 \Rightarrow c_{\text{rdu}}(E) = 1$ (fixed endpoints)
3. $p(E) \geq p(F) \Leftrightarrow c_{\text{rdu}}(E) \geq c_{\text{rdu}}(F)$ (order-preserving)

It follows that c_{rdu} is a *Choquet capacity* (see Choquet 1954), satisfying normalization, non-negativity, and a monotonicity condition: $E \subseteq F$ implies $c_{\text{rdu}}(E) \leq c_{\text{rdu}}(F)$. It need not, however, be additive.[2]

Next, we understand a *risk function* (r) in the usual way as a (continuous) function from $[0,1]$ back into $[0,1]$ satisfying:

1. $r(0) = 0$ and $r(1) = 1$ (boundary normalization)
2. $x < y$ implies $r(x) < r(y)$ (strictly increasing)

Furthermore, let r_0 designate the 'neutral' risk function such that $r_0(x) = x$ for all $x \in [0,1]$.

[2] Nothing hangs on transforming p into some non-additive capacity. Where $\mathcal{S}$ is finite, there are cases where preferences maximize expected utility relative to exactly one probability function even though there are other probability functions with which it is ordinally equivalent. As such, the present example could be adjusted such that c_{rdu} is both additive and a non-trivial transformation of p. This would bring it more in line with standard rank-dependent utility theories, where the credence function is typically assumed to be additive.

Finally, let the *rank-dependent utility valuation* (*rdu*) be defined relative to a choice of c, r and u such that, if $\langle E_1, \ldots, E_n \rangle$ is the most coarse-grained ordered partition of $\mathcal{S}$ where $f_{E_i} = x_i$ implies $u(x_1) \geq \ldots \geq u(x_n)$, then

$$rdu^{(c,r,u)}(f) = \sum_{i=1}^{n-1} r\left(c\left(\bigcup_{j=1}^{i} E_j\right)\right)(u(x_i) - u(x_{i+1})) + u(x_n).$$

Now note two things. First, if c is additive, then *rdu* and *eu* coincide under the neutral risk function—i.e., $rdu^{(c,r_0,u)}$ just is $eu^{(c,u)}$ whenever c is additive. Second, because c_{rdu} is obtained from p by a non-trivial order-preserving transformation, we can always choose a risk function $r_{\text{rdu}} \neq r_0$ such that

$$r_{\text{rdu}}(c_{\text{rdu}}(E)) = p(E) = r_0(p(E)).$$

Those facts together imply that

$$rdu^{(c_{\text{rdu}},r_{\text{rdu}},u)} = rdu^{(p,r_0,u)} = eu^{(p,u)}.$$

Consequently, (MEU) holds just in case

$$f \succsim g \Leftrightarrow rdu^{(c_{\text{rdu}},r_{\text{rdu}},u)}(f) \geq rdu^{(c_{\text{rdu}},r_{\text{rdu}},u)}(g). \quad \text{(RDU)}$$

So Bayesia's preferences maximize *eu* relative to p and u if and only if they maximize *rdu* relative to c_{rdu}, r_{rdu} and u.

2.5 State-dependent utilities

The final example relates directly to the problem of state-dependent utilities, as discussed by Seidenfeld et al. (1990), Schervish et al. (1990), Karni (1993) and Baccelli (2020).

Take a *state-dependent utility function* (v) to be a mapping from consequence-state pairs to real numbers; a *state-dependent utility valuation* (*sdu*) is then defined relative to c and v such that

$$sdu^{(c,v)}(f) = \sum_{s \in \mathcal{S}} c(s)v(f(s), s).$$

Next, let $c_{\text{sdu}} \neq p$ be *any* probability function such that $c_{\text{sdu}}(s) > 0$ whenever $p(s) > 0$. Observe that p and c_{sdu} may have very little in common. In particular, c_{sdu} need not be related to p by any order-preserving transformation, nor even by any *bijective* transformation. There may be some E, F where $p(E) > p(F)$ but not $c_{\text{sdu}}(E) > c_{\text{sdu}}(F)$ (or vice versa), and even some where $p(E) = p(F)$ but not $c_{\text{sdu}}(E) = c_{\text{sdu}}(F)$ (or vice versa).

Nevertheless, suppose v_{sdu} is given by

$$v_{\text{sdu}}(x, s) = \begin{cases} \frac{p(s)}{c_{\text{sdu}}(s)} u(x) & \text{if } p(s) > 0 \\ \text{anything} & \text{otherwise.} \end{cases}$$

Now whenever $p(s) > 0$,

$$p(s) = c_{\text{sdu}}(s)\frac{p(s)}{c_{\text{sdu}}(s)}.$$

As such, differences between $p(s)$ and $c_{\text{sdu}}(s)$ can be absorbed into the differences between $u(x)$ and $v_{\text{sdu}}(x,s)$:

$$\begin{aligned} eu^{(p,u)}(f) &= \sum_{s\in\mathcal{S}} p(s)u(f(s)) \\ &= \sum_{\substack{s\in\mathcal{S}\\ p(s)>0}} c_{\text{sdu}}(s)\frac{p(s)}{c_{\text{sdu}}(s)}u(f(s)) \\ &= \sum_{s\in\mathcal{S}} c_{\text{sdu}}(s)v_{\text{sdu}}(f(s),s) \\ &= sdu^{(c_{\text{sdu}},v_{\text{sdu}})}(f). \end{aligned}$$

So (MEU) holds just in case

$$f \succsim g \;\Leftrightarrow\; sdu^{(c_{\text{sdu}},v_{\text{sdu}})}(f) \geq sdu^{(c_{\text{sdu}},v_{\text{sdu}})}(g). \tag{SDU}$$

That is, Bayesia's preferences maximize *eu* relative to p and u if and only if they maximize *sdu* relative to c_{sdu} and v_{sdu}.

3. Three species of realism

We have seen six 'representational packages', each corresponding to a biconditional that holds just in case all the others do. The question now is which, if any, are really just different ways of representing the same thing.

It's clear, I think, what the constructivist will say: the inequality on the right of (MEU) is nothing more than a fancy way of representing the preferences on the left, and the inequalities in (ADD), (AFF), (MUL), (RDU) and (SDU) all represent the same preferences—so there's no meaningful difference between them. Realism, on the other hand, is more complicated. The minimal commitment to the idea that p represents *something* psychologically real doesn't yet tell us what that psychological reality *is*, and therefore which properties of the probability function are meaningful as opposed to mere artifacts of convention.

	Representationally equivalent packages
numerical realism	(MEU)
scalar realism	(MEU), (ADD)
structural realism	(MEU), (ADD), (AFF), (MUL)
constructivism	(MEU), (ADD), (AFF), (MUL), (RDU), (SDU)

In what follows, I will therefore outline three species of realism: *numerical*, *scalar* and *structural*. As summarized in the table, they are progressively more permissive views about which biconditionals say the same thing as (MEU). But do note that the numerical/scalar/structural trichotomy is not meant to be exhaustive. For instance, Joyce (2015: 416–19) can be read as suggesting *affine* *realism*, according to which any linear transformation of p represents the same credences as p. The purpose is not to describe every conceivable species of realism, but to highlight a natural progression of ideas leading to the view I take to be most plausible—to wit, structural realism.

3.1 Numerical realism

Numerical realism can be characterized simply as the view that there is at most one correct way to numerically represent an agent's credences—so if any credence function accurately represents an agent's credences, it does so *uniquely*. Consequently, numerical realists will interpret all six biconditionals as corresponding to mutually exclusive psychological models.

Those encountering decision theory for the first time often seem to adopt an interpretation along these lines. Students will, for instance, often presume that the numbers assigned by a subjective probability function should correspond more or less directly to how the agent represents their own uncertainty—such as when they're plugging values into a decision matrix, calculating odds, or introspecting on their degrees of confidence. After all, if Bayesia *does* have particular numerical values explicitly in mind, then it's natural to think that an adequate numerical representation of her credences should reflect those exact values. Anything else would fail to represent what's *really* going on inside her head.

More generally, numerical realism often seems to be treated as the default interpretation of subjective probabilities in decision theory. Consider the following from Alan Hájek, responding to Zynda's original example:

> According to probabilism, rationality requires an agent's *credences* to obey the probability calculus. We have rival ways of representing an agent whose preferences [maximize expected utility relative to p and u]; which of these representations [i.e., p or c_{aff}] correspond to her *credences*? In particular, why should we privilege the probabilistic representation? (2008: 806)

And a little later:

> It might be objected that the 'rival' representations are not *really* rival. Rather, the objection goes, they form a family of isomorphic representations, and choosing among them is merely a matter of convention… First, a perhaps flat-footed reply: I understand 'probabilism' to be *defined* via Kolmogorov's axiomatization of probability. So, for example, a non-additive measure is not a probability function. (2008: 806)

Flat-footed or not, it's hard to make sense of the reply as anything but a non sequitur unless numerical realism is being tacitly assumed.

That might need more explanation. First, while probabilism is often formulated as the view that 'rationality requires an agent's *credences* to obey the probability calculus', in a literal sense that *cannot* be right. What obeys the probability calculus are real-valued functions, and credences are not real-valued functions. Probably, I would say, they are patterns of synaptic interconnection identified via their normal role in relation to evidence and behavior. But *whatever* they may be, credences and credence functions are not the same thing on any sensible version of realism.

So the usual formulation of probabilism cannot be taken literally. More plausibly, probabilism is a constraint on credences expressed in terms of their representation. At first pass, it says that rational credences must be structured so they can be accurately represented by some probability function.[3] However, on *that* understanding there is no immediate conflict between probabilism and non-probabilistic representations. After all, we represent lots of things in different ways, so there's nothing incoherent about saying that an agent's credences might be accurately represented by both p and c_{aff}. Thus, the mere fact that c_{aff} violates the probability axioms cannot, *by itself*, imply that c_{aff} cannot represent rational credences. To get that result, one needs the further premise that an agent's credences admit *at most one* correct numerical representation. Otherwise, it's unclear why it should matter that 'probabilism' is usually *defined* by Kolmogorov's axiomatization.

Hájek goes on to suggest that

> … one might want to have a broader understanding of 'probabilism', encompassing any transformation of a probability function and a correspondingly transformed combination rule for utility that yields the same ordinal representation of preferences. If that is what is intended, then probabilism should be stated in those terms, and not in the flat-footed way that is entirely standard. (2008: 806)

But the point is that we *don't* need to broaden the definition of probabilism! We agree that probability functions are strictly defined by the Kolmogorov axioms, and that probabilism is properly understood in terms of probability functions in this strict sense. The issue is whether probabilism, so understood, implies that c_{aff} cannot represent the same credences as p. It does not—not unless we assume that the probabilistic representation is the only accurate representation. But that's just numerical realism, or something to the same effect.

3.2 Scalar realism

Numerical realism as just defined concerns *credences*. The parallel view for utilities is not widely accepted, since most decision theorists take utilities to be measured on an interval scale, like temperatures in Celsius or Fahrenheit. What matters, therefore, is not the specific numerical values in any representation, but what each has in common with the others.

It's worth getting clear on *why* this is the orthodox view. Ultimately, it boils down to two facts:

[3] A more precise statement requires specifying the representational format in which a rational agent's credences should be represented probabilistically. I will say more about representational formats shortly; but for the present discussion the rough formulation will suffice.

(i) if $\succsim$ maximizes *eu* relative to c and u, and u' is any positive affine transformation of u, then $\succsim$ also maximizes *eu* relative to c and u'; and
(ii) the same is not true in general for anything other than positive affine transformations.

Let me put it heuristically like this. Suppose Bayesia explicitly represents her credences and utilities using p and u, and wants to calculate expected utilities. The natural way to do so is to *multiply-then-add*: for each act f, she multiplies the utilities of its consequences at each relevant event by the probabilities she assigns to those events, then adds the results. But then she quickly realizes that what matters is not the particular values she uses to represent her utilities, but the ratios of their differences. If u and u' are related by a positive affine transformation, then—holding p fixed—it makes no difference whether she multiplies-then-adds with one or the other. The lesson she draws from this is that as far as expected utility theory is concerned, there is no meaningful difference between u and u'.

The point is not that (MEU) is true only if Bayesia employs the multiply-then-add *procedure*. That's too strong. (MEU) characterizes a *relationship* between Bayesia's credences, utilities, and preferences; and the multiply-then-add procedure is just one way that relationship might come to be realized. The lesson is that there's no meaningful difference between u and u' within the context of expected utility theory. If both the probability function and the form of the valuation are held fixed, then u and u' play exactly the same role—so there's no reason *not* to interpret them as representing the same thing.

But if *that's* right, then the same reasoning implies that credences ought to be measured on a ratio scale. After all,

(i) if $\succsim$ maximizes *eu* relative to c and u, and c' is any positive scalar transformation of c, then $\succsim$ also maximizes *eu* relative to c' and u; and
(ii) the same is not generally true for anything other than positive scalar transformations.

Just as with utilities, then, Bayesia ought to recognize that what matters is not the particular numerical values chosen to represent her credences, but the ratios between them. Holding the utility function and valuation fixed, it makes no difference whether she multiplies-then-adds with p, with c_{add}, or with any other credence function related to p by a positive scalar. The differences between them are theoretically superfluous, and thus properly treated as meaningless.

The scalar realist accepts this conclusion—and rightly so. They recognize that realism *about credences* doesn't imply realism *about the numerical values used to represent them*. Better to see credence as a psychological quantity representable on a range of scales all related by the appropriate transformations. In this case, the difference between p and c_{add} is no more significant than the difference between measuring length in meters or inches—they are just different ways of stretching and squeezing the numbers while preserving the same relevant information. For the same reason, the scalar realist concludes that (MEU) and (ADD) are simple notational variants of one another, differing only by an arbitrary choice of scale.

However, the same line of reasoning does not carry over to any of the other examples in §2, since none of those involve credence functions related to p by a scalar transformation. One cannot simply swap p for c_{mul}, say, or for c_{rdu}, while holding *everything* else fixed and expect no differences in the preferences that result. Thus, the scalar realist concludes that (AFF), (MUL), (RDU), and (SDU) involve more than just a difference in scale, and therefore represent something different from (MEU).

3.3 Structural realism

Scalar realism is a halfway house between numerical realism and structural realism: it's more permissive than numerical realism, but for the latter it doesn't go far enough. The scalar and structural realists agree that credences are measurable on a ratio scale; so p represents Bayesia's credences only if c_{add} does too. It does not follow, however, that any credence function *not* related to p by a positive scalar transformation therefore *cannot* represent the same credences. The structural realist knows better, since they recognize that what's meaningful in any representation is only sensibly understood relative to a choice of representational format.

Consider length. Let m be the function from objects to numbers corresponding to the meter scale. Then m has the following properties in common with all conventional measures of length:

1. m assigns each object's length a unique non-negative real number.
2. An object a is at least as long as b just in case $m(a) \geq m(b)$.
3. If a is partitioned into two parts b and c by a plane orthogonal to the dimension along which their lengths are measured, then $m(a) = m(b) + m(c)$.

Together these properties characterize the *additive format*. The first part fixes the range of values, the second fixes the ordering convention, and the third gives the format its additive flavor. Moreover, any other representation of length *in the same format* will relate to m via some positive scalar transformation. Since those transformations preserve ratios, we therefore say length is measured on a ratio scale.

It has long been recognized, however, that we might equally well represent the relation between the length of a whole and the lengths of its parts using multiplication. (See Hölder 1901; Krantz et al. 1971: 99–102.) The additive format is encoded by $\langle \mathbb{R}^{\geq 0}, \geq, + \rangle$: lengths are represented by non-negative reals, length-ordering by the usual ordering of those reals, and the part-whole relation by addition. But for any $\alpha > 1$, the map $x \mapsto \alpha^x$ is an (order-preserving and operation-preserving) isomorphism from $\langle \mathbb{R}^{\geq 0}, \geq, + \rangle$ to $\langle \mathbb{R}^{\geq 1}, \geq, \times \rangle$. Consequently, any representation of length in the additive format can be 'converted' into a scale on $[1, \infty)$ in which the part-whole relation is captured by multiplication rather than addition. That is, if m is the meter scale as above, then $m_{\text{mul}} = \alpha^m$ with $\alpha > 1$ is a *multiplicative scale* such that

$$m_{\text{mul}}(a) = m_{\text{mul}}(b) \cdot m_{\text{mul}}(c).$$

And there's nothing special about multiplication either. For any constant β, let $+^\beta$ be the β-*affine addition* operation,

$$x +^\beta y \coloneqq x + y - \beta.$$

We can then obtain an isomorphism from $\langle \mathbb{R}^{\geq 0}, \geq, + \rangle$ to $\langle \mathbb{R}^{\geq \beta}, \geq, +^\beta \rangle$ via $x \mapsto \alpha x + \beta$ for $\alpha > 0$. Accordingly, $m_{\text{aff}} = \alpha m + \beta$ will be a β-affine scale such that

$$m_{\text{aff}}(a) = m_{\text{aff}}(b) + m_{\text{aff}}(c) - \beta.$$

Indeed, for *any* bijection $\varphi : \mathbb{R} \to \mathbb{R}$ we can define a corresponding format. Let

1. $\mathbb{R}^\star = \varphi[\mathbb{R}^{\geq 0}]$,
2. $xRy \Leftrightarrow \varphi^{-1}(x) \geq \varphi^{-1}(y)$,
3. $x \oplus y \coloneqq \varphi(\varphi^{-1}(x) + \varphi^{-1}(y))$.

Then $\langle \mathbb{R}^{\geq 0}, \geq, + \rangle$ is isomorphic to $\langle \mathbb{R}^\star, R, \oplus \rangle$, in the relevant sense. Thus we can define an $\oplus$-ive representation $m_\oplus = \varphi \circ m$ under which

1. $m_\oplus$ assigns each object's length a value in $\mathbb{R}^\star$.
2. And object a is at least as long as b just in case $m_\oplus(a) R m_\oplus(b)$.
3. if a is composed lengthwise of disjoint parts b, c, then $m_\oplus(a) = m_\oplus(b) \oplus m_\oplus(c)$.

Perhaps the most crucial thing to recognize about converting between representational formats is that how we express relations between quantities is sensitive to how those quantities are represented. Consider Hooke's Law, which describes how elastic materials deform proportionately to the force applied when stretching. It is standardly expressed $F = kx$, where the stretching force F is measured in Newtons, k is the material's 'stiffness' in Newtons per meter, and x is the distance in meters from its resting position. In the present notation, x is just $m(d)$, where d is the length of that stretch and m is the meter scale. Thus the same law can be written

$$F = k \cdot m(d).$$

But suppose we represent length on a multiplicative scale instead, keeping all else fixed. In that case, simply substituting $m_{\text{mul}}(d)$ for $m(d)$ breaks the law—for instance, the equation assigns a nonzero force ($F = k$) to the resting position $m_{\text{mul}}(d) = 1$ (i.e., $m(d) = 0$). Instead, a proper reconstruction requires changing the *form* of the law. The neatest version uses F_{mul}, which is the multiplicative reformatting of the Newton scale by the same transformation $x \mapsto \alpha^x$ used to go from m to m_{mul}. Thus:

$$F_{\text{mul}} = m_{\text{mul}}(d)^k.$$

The thing to note here is that the underlying *physical relationships* being represented are the same for either version of the law—all that's changed is how we *express* them. It's the same structure, just a different format.

In a nutshell, structural realism is just the application of lessons to the case of credences. First, to say that p represents Bayesia's credences on a ratio scale does not imply that all legitimate representations of those same credences must be related to p by some positive scalar transformation—that's true only if we stick to the usual additive representational format. Second, when converting between formats, we should not hold all else fixed. The very same relationships expressed via the expected utility valuation can be equally well expressed via the multiplicative and affine valuations, which are just the natural re-expressions of the expected utility valuation relative to the conversion from additive to multiplicative and affine-additive formats respectively. Thus, the structural realist concludes, (ADD), (AFF) and (MUL) all express the same thing as (MEU)—because they're all just different ways of expressing the same structures.

4. Structural realism and isomorphism

That's the basic idea, but 'same structure' still needs to be made precise. The challenge is to give an account that's permissive enough to count (AFF) and (MUL) as equivalent to (MEU), and more generally to allow for *any* possible variation in representational format, but not *so* permissive that 'same structure' just reduces to 'same preferences'—otherwise, structural realism would be indistinguishable from constructivism. In this section, I provide such an account and then apply it to the examples in §2, with the result that (ADD), (AFF), and (MUL) all turn out to be re-expressions of (MEU), while (RDU) and (SDU) express something meaningfully different.

4.1 Theories, models, and isomorphism

The first step is to pin down the objects to be compared. What I will henceforth call a *decision theory* is characterized by

(i) a *valuation*, which assigns values to acts relative to some tuple of parameters;
(ii) a *decision rule*, which establishes an equivalence between preferences and the relative value of acts as determined by that valuation; and
(iii) a set of *admissibility constraints* on the relevant parameter tuples.

A *model* of that theory then consists of a preference ordering together with an admissible parameter tuple, such that the ordering satisfies the decision rule given those parameters.

Think of the theory as picking out a class of possible psychologies by reference to how preferences relate to credences and utilities via some valuation. A model then represents one possible instance of that psychological structure. For example, *expected utility theory* is characterized by the rule that preferences maximize expected utility relative to an admissible parameter pair, where the first is any real-valued function on events satisfying normalization, non-negativity and additivity, while the second is any real-valued function on consequences. An *expected utility model* is then a triple $\langle \succsim, c, u \rangle$ such that c is a probability function, u is a utility function, and $\succsim$ maximizes eu relative to c and u.

Note one: the admissibility constraints can serve several purposes. Some are required for the associated valuation to behave as intended. As noted earlier, for example, the eu valuation is not well-defined except relative to an additive credence function, so additivity is a minimal condition for expected utility theory (so defined) to make sense. Non-negativity, by contrast, is not implied by partition-independence, and more generally is not required for the *eu* valuation to be well-defined. Instead, it serves to rule out preference patterns that would otherwise violate intuitive dominance conditions—such as Savage's postulate P3, which roughly says that replacing a worse consequence with a better one should make an act better (or at least not worse).

Note two: I have intentionally characterized decision theories and their models in a very fine-grained way, in order to provide a neutral formal space in which different interpretations can be compared without prejudging questions of representational equivalence. It is not to be assumed, therefore, that different models within a theory represent distinct psychologies. If u and u' are related by a positive affine transformation, for instance, then two models $\langle \precsim, c, u \rangle$ and $\langle \precsim, c, u' \rangle$ presumably represent the same thing, given that u and u' presumably represent the same thing.

But those are questions about the equivalence of models *within* a decision theory, whereas our question concerns equivalence *across* decision theories. According to structural realism, this should be a matter of when the models are corresponding instances of the same general structure. With that in mind, consider expected utility theory. The relevant structure to be preserved is:

1. the *structure of preference determination*: how independent parameters combine to determine preferences; and
2. the *internal structure* of the parameters themselves.

Start with preference determination. Under a realist interpretation, expected utility theory (T_{eu}) posits a simple causal structure: credences and utilities are independent factors that together determine preferences via the expected utility rule. This independence is captured within the theory through recombinability. Say that a parameter a is *admissible* in a decision theory T just in case a occurs as a coordinate of some admissible parameter tuple in T. Then, if c and c' are any two credence functions admissible within T_{eu}, and u and u' are any two admissible utility functions, c can be paired with u, c with u', c' with u, and c' with u'. More generally, if A_i is the admissible parameter space for the i-th parameter, then T posits n independent parameters when the set of the theory's admissible parameter tuples A_T factors as

$$A_T = A_1 \times \cdots \times A_n.$$

But preserving this structure of preference determination is not enough, since the parameters themselves are not structureless objects. That is, if T is to have the 'same structure' as T_{eu}, then there ought to be a one-to-one correspondence ψ between the corresponding parameter spaces of the two theories where $\psi(a)$ can legitimately be seen as representing the same thing as a—albeit possibly in some very different format. And as we saw in §3.3, this essentially means that $\psi(a)$ must be related to a by some pointwise bijection φ such that

$$\psi(a)(x) = \varphi(a(x)).$$

The core idea, then, is that any re-expression of expected utility theory ought to preserve both independence and internal structure: there should be a one-to-one correspondence between the models of T_{eu} and those of the alternative theory T that preserves preference determination, maps independent credence and utility parameters to their independently specifiable counterparts in T, and is induced by structure-preserving maps between parameter spaces.

4.2 (ADD), (AFF) and (MUL)

The easy case is when two theories T_{a} and T_{b} have the same number of independent parameters, with corresponding parameters defined on the same domains. Let $A_1, \ldots, A_n$ be the admissible parameter spaces in T_{a}, and $B_1, \ldots, B_n$ the corresponding admissible parameter spaces in T_{b}. Where C is any parameter space, let $\text{Dom}(C)$ denote the common domain of its parameters. Under these circumstances, T_{a} is *isomorphic* to T_{b} just in case there are bijections $\psi_i : A_i \to B_i$ for each $i = 1, \ldots, n$ satisfying:

Pointwise correspondence. There is a bijection φ_i such that for all $a \in A_i$ and $x \in \text{Dom}(A_i)$, $\psi_i(a)(x) = \varphi_i(a(x))$.

Preference determination. For all $a_1 \in A_1, \ldots, a_n \in A_n$ and any preferences $\succsim$, $\langle \succsim, a_1, \ldots, a_n \rangle$ is a model of T_{a} if and only if $\langle \succsim, \psi_1(a_1), \ldots, \psi_n(a_n) \rangle$ is a model of T_{b}.

In other words, the models of T_{a} and T_{b} can be placed in a one-to-one correspondence induced by independent pointwise transformations of the corresponding parameters.

The multiplicative version of expected utility theory (T_{mul}) is a natural example. It is characterized by the rule that preferences maximize *mul* relative to a credence and utility function, with the main constraint that the credence function must satisfy renormalization, unit lower-boundedness and multiplicativity. A quick proof of the isomorphism:

Proof. Let $A_{\text{c}} \times A_{\text{u}}$ be the set of parameter tuples admissible in T_{eu}, and $B_{\text{c}} \times B_{\text{u}}$ the set admissible in T_{mul}. Let $\psi_{\text{c}} : A_{\text{c}} \to B_{\text{c}}$ be defined pointwise by $\psi_{\text{c}}(c)(E) = 10^{c(E)}$, and let $\psi_{\text{u}} : A_{\text{u}} \to B_{\text{u}}$ be the identity map id. Since the corresponding value-scale maps are $\varphi_{\text{c}}(x) = 10^x$ and $\varphi_{\text{u}} = \text{id}$, *pointwise correspondence* holds. Furthermore, for every admissible pair (c, u) in T_{eu}, the pair $(\psi_{\text{c}}(c), u)$ is admissible in T_{mul}, and the ordinal equivalence of $eu^{(c,u)}$ and $mul^{(\psi_{\text{c}}(c),u)}$ implies that $\langle \succsim, c, u \rangle$ is a model of T_{eu} if and only if $\langle \succsim, \psi_{\text{c}}(c), u \rangle$ is a model of T_{mul}. So *preference determination* also holds. So T_{eu} is isomorphic to T_{mul}. □

Notice two things about this. First, the maps ψ_{c} and ψ_{u} induce the desired model-level equivalence—in particular, $\langle \succsim, c_{\text{mul}}, u \rangle$ is the image of $\langle \succsim, p, u \rangle$ under ψ_{c} and ψ_{u}. Thus the structural realist interprets (MEU) and (MUL) as representing the same thing: *they describe corresponding models within isomorphic theories.* Second, ψ_{c} preserves the distinctive roles played by the probabilistic constraints in T_{eu}. Additivity, for example, is required for partition-independence in the *eu* valuation; and multiplicativity is required for partition-independence in the *mul* valuation. Likewise, unit lower-boundedness in T_{mul} is the direct theoretical counterpart of non-negativity in T_{eu}: it prevents credence functions in T_{mul} from assigning values on both sides of 1, which would otherwise permit P3-violating preferences under *mul*.

The point generalizes: *any* bijective transformation of p induces a corresponding decision theory isomorphic to expected utility theory. As mentioned in §3.3, if $\varphi : [0,1] \to \mathbb{R}^\star$ is any bijection, for some $\mathbb{R}^\star \subseteq \mathbb{R}$, then we can transport the structure of the standard probabilistic format along φ to obtain a corresponding format $\langle \mathbb{R}^\star, R, \oplus \rangle$. Then, if $c_\varphi = \varphi \circ p$, c_φ satisfies the φ-transformed analogues of the usual probabilistic constraints:

1. $c_\varphi(\mathcal{S}) = \varphi(1)$ (φ-normalization)
2. $c_\varphi(E) R i$ (i R-bounded)
3. $E \cap F = \varnothing$ implies $c_\varphi(E \cup F) = c_\varphi(E) \oplus c_\varphi(F)$ ($\oplus$-ivity)

A corresponding valuation can then be obtained simply by undoing the transformation inside the expected utility sum:

$$v_\varphi^{(c,u)}(f) = \sum_{E \in \mathcal{E}_f} \varphi^{-1}(c(E)) u(f_E).$$

As such,

$$v_\varphi^{(c_\varphi,u)}(f) = eu^{(p,u)}(f).$$

Finally, any order-preserving transformation of $v_\varphi^{(c,u)}$ will preserve the ordering of acts, and so yields yet another theory isomorphic to T_{eu}.

The (ADD), (AFF), and (MUL) examples are all just variations on the same general recipe: reformat the credence scale, adjust the valuation and admissibility constraints to fit, and then rewrite it all in a more convenient form if need be. Since the resulting theories are isomorphic to T_{eu}, the structural realist treats (ADD), (AFF), and (MUL) as notational variants of (MEU).

4.3 (RDU) and (SDU)

However, things are not always so easy. A model in expected utility theory can represent the same thing as a model in another decision theory even when the theories are not isomorphic. This can happen when the other theory includes expected utility theory—or something isomorphic to it—as a proper subtheory. Such cases can arise in a number of ways: by relaxing admissibility constraints (e.g., weakening additivity), altering the structure of parameters that feed into a valuation (e.g., shifting from state-invariant to state-dependent utilities), or adding new parameters entirely (e.g., a risk function).

Both (RDU) and (SDU) raise complications of this kind. That is, each corresponds to a model within a richer decision theory into which expected utility theory can be embedded as a special case. Dealing with such cases requires a more general notion of isomorphism that takes a bit of work to fully spell out, which can be found in the Appendix. For the present discussion, the rough ideas should suffice.

Start with (RDU). Let *rank-dependent utility theory* (T_{rdu}) for present purposes be characterized by the rule that preferences maximize *rdu* relative to three independent parameters—labeled $1, 2, 3$—subject to the constraints:

- parameter 1 is a capacity;
- parameter 2 is a risk function;
- parameter 3 is a utility function.

Since T_{rdu} has three parameters and T_{eu} has two, the earlier definition of isomorphism does not apply. But there's an obvious extension: instead of mapping every independent parameter in T_{eu} to a single parameter in T_{rdu}, they can instead be mapped to a tuple of one or more parameters.

Given that, the most obvious way to interpret rank-dependent utility theory as a generalization of expected utility theory is to treat the latter as the special case of the former whereby (i) the agent's credences are represented by a probability function, and (ii) their risk attitudes are a psychological spinning wheel that contribute nothing to the determination of preferences. Call this the *standard interpretation*; it corresponds to the isomorphism of T_{eu} with a subtheory of T_{rdu} established by

(i) $\psi_{\text{c}}(c) = (c, r_0)$, where r_0 is the neutral risk function; and
(ii) $\psi_{\text{u}} = \text{id}$.

In this case, $\langle \succsim, c_{\text{rdu}}, r_{\text{rdu}}, u \rangle$ is plainly not the image of $\langle \succsim, p, u \rangle$. So, under the standard interpretation, (RDU) and (MEU) represent very different psychologies—i.e., one where Bayesia has substantive risk attitudes, the other where she's risk neutral.

But there is another way to see things. On the standard interpretation, parameter 1 represents the agent's credences, while parameter 2 represents a separate psychological factor that happens to be a spinning wheel in expected utility cases. However, one might instead interpret parameters 1 and 2 as a decomposition of what expected utility theory represents as a single parameter—that is, if $r \circ c' = p$, then the pair (c', r) is one way of instantiating the credences represented in T_{eu} by p. Call this the *decompositional interpretation*; it corresponds to the family of isomorphisms from T_{eu} into subtheories of T_{rdu}, established by

(i) $\psi_{\text{c}}(c) = (c', r)$ such that $c'(E) = r^{-1}(c(E))$ for all E; and
(ii) $\psi_{\text{u}} = \text{id}$.

And in this case, there a actually several ways to embed T_{eu} into T_{rdu} under which $\langle \succsim, c_{\text{rdu}}, r_{\text{rdu}}, u \rangle$ is the image of $\langle \succsim, p, u \rangle$. But that doesn't mean we should read (RDU) as equivalent in meaning to (MEU). On the decompositional interpretation, it's the *combination* of c_{rdu} and r_{rdu} that represents *one particular way* of instantiating the credences represented more coarsely by p. Better to say, then, that (RDU) expresses a fine-grained realization of (MEU)—one among many—and that c_{rdu} by itself represents only one component of what p is supposed to represent in (MEU).[4]

[4] Because c_{rdu} is a bijective transformation of p, there must be *some* theory within which c_{rdu} represents the same credences as p does in T_{eu}. But that's not the same question as whether c_{rdu} *as it appears in* (RDU) represents the same credences as p *as it appears in* (MEU).

Turn now to (SDU). Let *state-dependent utility theory* (T_{sdu}) be characterized by the rule that preferences maximize *sdu* relative to two independent parameters, where

- parameter 1 is a probability function;
- parameter 2 is a state-dependent utility function.

Though T_{sdu} is clearly similar to T_{eu}, the two are far from isomorphic—in particular, the additional structure of state-dependent utility functions permits a much wider range of preference patterns. For a simple illustration, suppose there are two states s_1, s_2 and two consequences x, y. Let f_{xy} be the act that yields x in s_1 and y in s_2, and define f_{xx}, f_{yy} and f_{yx} similarly. Then T_{sdu} can represent preferences such as

$$f_{xy} \succ f_{xx} \sim f_{yy} \succ f_{yx}.$$

This can happen if x is better than y in state s_1 but worse than y in state s_2. But these preferences violate Savage's P3, so cannot arise within T_{eu}.

Still, there's an obvious way to see T_{eu} as 'contained in' T_{sdu}: simply identify every utility function in T_{eu} with the state-dependent utility function that assigns the same value to a consequence in every state. More precisely, let $\widehat{u}$ denote the state-dependent utility function such that $\widehat{u}(x, s) = \widehat{u}(x, s') = u(x)$ for all s, s' and x; and let $T^{\star}_{\text{sdu}}$ be the subtheory of T_{sdu} whose admissible v are exactly these 'state-constant' functions. Then $T^{\star}_{\text{sdu}}$ is isomorphic to T_{eu}: if $\psi_{\text{c}} = \text{id}$ and $\psi_{\text{u}}(u) = \widehat{u}$, then $\langle \succsim, c, u\rangle$ is a model of T_{eu} just in case $\langle \succsim, c, \widehat{u}\rangle$ is a model of $T^{\star}_{\text{sdu}}$. Observe, however, that in this case $\langle \succsim, c_{\text{sdu}}, v_{\text{sdu}}\rangle$ is not the image of $\langle \succsim, p, u\rangle$. That is, the model to which (SDU) corresponds lies elsewhere within state-dependent utility theory, outside the part to which T_{eu} most naturally corresponds—so (SDU) doesn't represent the same thing as (MEU).

Could there be some other embedding of T_{eu} into T_{sdu}, one where $\langle \succsim, c_{\text{sdu}}, v_{\text{sdu}}\rangle$ is the image of $\langle \succsim, p, u\rangle$? Short answer: no. (Slightly longer answer: not relative to the natural domain-composition map; see the Appendix.) If c_{sdu} and p are any two distinct probability functions related by some pointwise bijection, then the ratios $p(s)/c_{\text{sdu}}(s)$ cannot be constant across all states where they are defined. So pick s, s' such that

$$\frac{p(s)}{c_{\text{sdu}}(s)} \neq \frac{p(s')}{c_{\text{sdu}}(s')}.$$

Since u is non-constant, pick x such that $u(x) \neq 0$. It follows that

$$v_{\text{sdu}}(x, s) = \frac{p(s)}{c_{\text{sdu}}(s)} u(x) \neq \frac{p(s')}{c_{\text{sdu}}(s')} u(x) = v_{\text{sdu}}(x, s').$$

But then there cannot be any function φ_{u} such that $\varphi_{\text{u}}(u(x))$ determines both $v_{\text{sdu}}(x, s)$ and $v_{\text{sdu}}(x, s')$. That is: even if c_{sdu} is a pointwise transformation of p, v_{sdu} cannot be a pointwise transformation of u. So $\langle \succsim, c_{\text{sdu}}, v_{\text{sdu}}\rangle$ cannot be the image of $\langle \succsim, p, u\rangle$ under any isomorphic embedding of T_{eu} into T_{sdu}.

5. Weak realism

I want to conclude by considering the relationship between structural realism and Zynda's *weak realism*. The two are close in spirit—both allow that the credences represented by p might also be represented by some non-additive function, provided the appropriate adjustments are made to the corresponding valuation. But they differ in key respects, and getting clear on these will help characterize what structural realism is not.

As Zynda describes it, weak realism says: first, look at all credence functions c that could combine somehow with some values or other to determine the same preferences $\succsim$ as p and u under eu; then, since p and c represent the same credences, what's 'real' in either is at most what they share:

> One might point out that [c_{aff}] is simply a linear transformation of p, and argue that in the case of probabilities (like utilities and temperatures) this is a difference that makes no difference. This approach commits… to taking as real properties of degrees of belief at most those properties that are common to *both* [p and c_{aff}]. (2000: 64)

Moreover, this *common-properties principle* needs to be applied consistently:

> To pursue this strategy consistently, one would have to investigate the properties of all possible [credence functions] that can, when combined in some way with some representation of value, produce a function that is some order-preserving transformation of [the *eu* valuation], since [c_{aff}] is only *one example* of an alternative *quantitative* [credence function] that can be formulated consistently with the axioms of expected utility theory. (2000: 64)

Zynda then argues that while these measures won't all have properties like additivity in common, they all correspond to the same relative confidence ordering over events. As such, he concludes:

> The concept of degree of belief on this strategy becomes a purely ordinal notion… (2000: 65)

The result is that weak realism is now closely associated with the idea that all meaningful information in the probabilistic representation of credences reduces to 'purely ordinal' information (e.g., Meacham & Weisberg 2011; Stefánsson 2018; Konek 2019; Neth 2025).

There are a few things to note here. First, 'can be combined in some way with some representation of value' does not adequately distinguish the (ADD), (AFF), and (MUL) examples from those like (RDU) and (SDU)—indeed, it suggests that any way of recovering the same preferences should count as a re-expression of (MEU). In this respect, structural realism goes beyond weak realism by supplying a criterion beyond mere preference preservation.

A deeper difference concerns the common-properties principle. First a clarification: the question is not whether the properties that p has in common with c_{aff} are *real*. Credences are not credence functions, so properties of the latter are not sensible candidates for psychological reality. The question, rather, is which of p's properties meaningfully *correspond* to something real; and the answer is *not* those properties that p has in common with every other credence function representing the same credences. Compare again the case of length: additive relations are not shared between the meter scale m and its multiplicative variant m_{mul}, but that doesn't mean addition is meaningless *in* m—it means only that meaningfulness is format-relative. Likewise for credences: if p and c_{aff} do not belong to the same representational format, then commonality between them is not a test of what's meaningful (or 'real') in either.

For the same reason, structural realism does not implicate a 'purely ordinal' conception of credences, simply because there may be more to the relevant structure than ordinal structure. In particular, structural realism is consistent with the view that genuine psychological structure can be found in how credence trades off with utility in the determination of preference, leading to meaningful differences between credence functions with the same ordinal information. But since this has been discussed elsewhere (e.g., Elliott 2024), I won't pursue it further here. The important point is just that 'same structure' need not reduce to 'same ordinal structure'. In that sense, structural realism does not suggest a 'purely ordinal' conception of credences.[5]

Finally, and most importantly, structural realism differs from weak realism in its ontology. Zynda (2000: 67) combines weak realism with Patrick Maher's view that 'attributions of probability and utility [are] essentially a device for interpreting a person's preferences' (Maher 1993: 9). On this point I firmly disagree. According to structural realism, probability and utility functions represent genuine psychological quantities independent of, and causally prior to, preferences. They are not just devices for interpreting preferences, and there is more to having credences representable by a probability function p than simply having preferences that maximize expected utility relative to p. It is, in other words, realism in the fullest sense; the difference between it and other forms of realism is just the recognition that one and the same psychological structure can be represented in drastically different ways.

References

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

Baccelli, J. (2020). The Problem of State-Dependent Utilities: A Reappraisal. *British Journal for the Philosophy of Science 72*(2), 617–634.

Buchak, L. (2014). Risk and Tradeoffs. *Erkenntnis 79*(6 Supplement), 1091–1117.

Choquet, G. (1954). Theory of capacities. *Annales de l'institut Fourier 5*, 131–295.

Elliott, E. (2024). *The Measurement of Subjective Probability.* Cambridge University Press.

[5] Zynda never says that *only* ordinal structure matters, and perhaps what he meant by 'purely ordinal' is something more like 'qualitative' (i.e., not restricted to ordinal information). So it's possible he did not intend weak realism to be committed to the 'purely ordinal' view. With that said, I am more concerned with how weak realism has been received in the literature, and for better or worse it is generally associated with that view.

Gul, F. (1992). Savage's Theorem with a Finite Number of States. *Journal of Economic Theory 57*, 99–110.

Hölder, O. (1901). Die Axiome der Quantitat und die Lehre vom Mass. *Ver. Verh. Kgl. Saschisis. Ges. Wiss. Leipzig, Math.-Physc. Classe 53*, 1–63.

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

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

Karni, E. (1993). Subjective Expected Utility Theory with State-Dependent Preferences. *Journal of Economic Theory 60*, 428–438.

Konek, J. (2019). Comparative Probabilities. In *The Open Handbook of Formal Epistemology*, pp. 267–348. PhilPapers Foundation.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971). *Foundations of Measurement, Vol. I: Additive and polynomial representations*. Academic Press.

Maher, P. (1993). *Betting on Theories*. Cambridge University Press.

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

Neth, S. (2025). Better Foundations for Subjective Probability. *Australasian Journal of Philosophy*, 1–22.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization 3*(4), 323–343.

Savage, L. J. (1954). *The Foundations of Statistics*. Dover.

Schervish, M. J., T. Seidenfeld, and J. B. Kadane (1990). State-dependent utilities. *Journal of the American Statistical Association 85*(411), 840–847.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica 57*(3), 571–587.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (1990). When Fair Betting Odds Are Not Degrees of Belief. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1: Contributed Papers*, 517–524.

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

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

Appendix

Suppose T_a (the *source*) has $n \geq 1$ independent parameters with admissible spaces $A_1, \ldots, A_n$, while T_b (the *target*) has $m \geq n$ parameters with admissible spaces $B_1, \ldots, B_m$. For simplicity, assume parameters in the same space share a domain, and all parameters in both theories have range V. To handle cases where T_a has fewer parameters than T_b or corresponding parameters have different domains, *pointwise correspondence* must be generalized. This involves three ideas:

1. For each $i = 1, \ldots, n$, a *parameter map* ψ_i sends each $a \in A_i$ to a tuple of target-side parameters.
2. A *domain-composition map* H_i gives the relevant source-target domain correspondence.
3. Relative to H_i, a *value-composition map* F_i says how the target-side parameters combine 'pointwise' to determine a pointwise transformation of the source-side parameter.

More precisely, for any non-repeating sequence $s = (C_1, \ldots, C_k)$ of parameter spaces, let a *block over* s be a subset $\mathfrak{B}$ of $C_1 \times \cdots \times C_k$; a *sub-block* of $\mathfrak{B}$ is a block over any non-empty subsequence of s. Recall that $\mathrm{Dom}(C)$ denotes the common domain of the parameters in C. If $\mathfrak{B}$ is a block over $C_1, \ldots, C_k$, let

$$\mathrm{Dom}(\mathfrak{B}) = \mathrm{Dom}(C_1) \times \cdots \times \mathrm{Dom}(C_k).$$

Thus, $\mathfrak{B}$ contains tuples of parameters, and $\mathrm{Dom}(\mathfrak{B})$ the corresponding tuples of arguments. Given this, a parameter map ψ_i is a bijection between A_i and some block $\mathfrak{B}_i$ formed from $B_1, \ldots, B_m$; and a domain-composition map H_i is an irreducible surjection from $\mathrm{Dom}(\mathfrak{B}_i^\star)$ to $\mathrm{Dom}(A_i)$, for some sub-block $\mathfrak{B}_i^\star$ of $\mathfrak{B}_i$. Here, if

$$H : \mathrm{Dom}(C_1) \times \cdots \times \mathrm{Dom}(C_k) \to Y,$$

then H is *irreducible* iff, for each $j = 1, \ldots, k$, some $x_j, x_j' \in \mathrm{Dom}(C_j)$ and fixed $x_\ell \in \mathrm{Dom}(C_\ell)$ for each $\ell \neq j$ satisfy

$$H(x_1, \ldots, x_j, \ldots, x_k) \neq H(x_1, \ldots, x_j', \ldots, x_k).$$

In other words, H_i is defined on the sub-block of $\mathfrak{B}_i$ formed from target-side parameter spaces whose arguments help determine the corresponding source-side argument. (Parameters outside $\mathfrak{B}_i^\star$ still matter for the value-composition map, but supply no arguments to H_i.)

For the value-composition map, we focus on a representative case that covers both (RDU) and (SDU). Suppose T_a has one parameter and T_b has two, where for all $a \in A$, $b \in B_1$ and $\rho \in B_2$,

$$a : \mathrm{Dom}(A) \to V, \qquad b : X \times Y \to V, \qquad \rho : Z \to V.$$

Now let $T_\mathrm{b}^\star$ be a subtheory of T_b whose admissible parameter tuples are exactly the members of some block $\mathfrak{B} \subseteq B_1 \times B_2$, and let H be such that

$$H : \mathrm{Dom}(B_1) \to \mathrm{Dom}(A).$$

Then T_{a} is isomorphic to $T^{\star}_{\mathrm{b}}$ relative to H just in case there's a bijection $\psi : A \to \mathfrak{B}$ satisfying:

Generalized pointwise correspondence. There is a bijection $\varphi : V \to V$ and a fixed map F such that, for all $a \in A$ and $(x, y) \in X \times Y$, if $\psi(a) = (b, \rho)$, then $F\big(b(x, y), \rho\big) = \varphi\big(a(H(x, y))\big)$.

Preference determination. For all $a \in A$ and preferences $\succsim$, $\langle \succsim, a\rangle$ is a model of T_{a} iff $\langle \succsim, \psi(a)\rangle^{\flat}$ is a model of $T^{\star}_{\mathrm{b}}$ (where $^{\flat}$ denotes the canonical flattening of nested tuples).

In other words, H matches up the arguments, while F matches up the values: it tells us how the target-side information at (x, y), possibly together with the additional parameter ρ, recovers the corresponding value of a.

The above can be generalized to theories with more parameters, arbitrary parameter domains, and different choices of H. But the notation quickly becomes intolerable, and the application to (RDU) and (SDU) in §4.3 doesn't require all the gory details. For (RDU), the isomorphisms between T_{eu} and subtheories of T_{rdu} are established on the credence side by:

1. $\psi_{\mathrm{c}} : A_{\mathrm{c}} \to \mathfrak{B}_{\mathrm{c}}$, where $\mathfrak{B}_{\mathrm{c}} \subset B_1 \times B_2$, B_1 is the set of capacities, and B_2 the set of risk functions.
2. $H_{\mathrm{c}} : \mathrm{Dom}(B_1) \to \mathrm{Dom}(A_{\mathrm{c}})$, with $H_{\mathrm{c}}(E) = E$.
3. $F_{\mathrm{c}} : [0, 1] \times B_2 \to [0, 1]$, with $F_{\mathrm{c}}(x, r) = r(x)$.
4. $\varphi_{\mathrm{c}} = \mathrm{id}$.

Hence $\psi_{\mathrm{c}}(c) = (c', r)$ implies $F_{\mathrm{c}}\big(c'(E), r\big) = \varphi_{\mathrm{c}}\big(c(H_{\mathrm{c}}(E))\big)$, or $r\big(c'(E)\big) = c(E)$.
For (SDU), T_{eu} is isomorphic to a subtheory of T_{sdu}. On the utility side:

1. $\psi_{\mathrm{u}} : A_{\mathrm{u}} \to \mathfrak{B}_{\mathrm{u}}$, where $\mathfrak{B}_{\mathrm{u}} \subset B_2$ and B_2 is the set of state-dependent utility functions.
2. $H_{\mathrm{u}} : \mathrm{Dom}(B_2) \to \mathrm{Dom}(A_{\mathrm{u}})$, with $H_{\mathrm{u}}(x, s) = x$.
3. $F_{\mathrm{u}} : \mathbb{R} \to \mathbb{R}$, with $F_{\mathrm{u}}(y) = y$.
4. $\varphi_{\mathrm{u}} = \mathrm{id}$.

Hence $\psi_{\mathrm{u}}(u) = v$ implies $F_{\mathrm{u}}\big(v(x, s)\big) = \varphi_{\mathrm{u}}\big(u(H_{\mathrm{u}}(x, s))\big)$, or $v(x, s) = u(x)$.

These isomorphisms are *relative to* the indicated domain-composition maps. Think of H as implicitly characterizing a *theory* of how a source-side domain is decomposed. I leave H formally unconstrained to avoid prejudging legitimate decompositions. Alternative domain-composition maps may therefore be possible, including artificial maps that encode the source domain in unexpected ways. For example, in the (SDU) case, $H_{\mathrm{c}} : \mathrm{Dom}(B_2) \to \mathrm{Dom}(A_{\mathrm{c}})$ would 'decompose' events as state-consequence pairs. Such maps could support different isomorphisms. But if a proposed domain-composition map is nonsense, the resulting embedding is just a formal curiosity. The examples above use the natural choices of H, and it's hard to imagine any sensible alternatives.