White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/168915/

Version: Accepted Version

Article:
Elliott, E (2022) What Is 'Real' in Interpersonal Comparisons of Confidence. Australasian Journal of Philosophy, 100 (1). pp. 102-116. ISSN: 0004-8402

https://doi.org/10.1080/00048402.2020.1849326

© 2020 Australasian Journal of Philosophy. This is an author produced version of an article published in Australasian Journal of Philosophy. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

WHAT IS 'REAL' IN INTERPERSONAL COMPARISONS OF CONFIDENCE

Edward Elliott

Abstract: According to comparativism, comparative confidence is more fundamental than absolute confidence. In a pair of recent papers, Stefánsson has argued that comparativism is capable of explaining interpersonal confidence comparisons. In this paper, I will argue that Stefansson's proposed explanation is inadequate; that we have good reasons to think that comparativism cannot handle interpersonal comparisons; and that the best explanation of interpersonal comparisons requires thinking about confidence in a fundamentally different way than that which comparativists propose—specifically, we should think of confidence as a dimensionless quantity.

Keywords: comparativism, confidence, interpersonal comparisons, degrees of belief, measurement

1. Introduction

Contrast two kinds of confidence states. On the one hand there's *comparative confidence*; this includes those states that we might attribute using, for example, 'is more confident that P than that Q' or 'is just as confident that P as that Q'. It is an essentially comparative attitude directed towards two (or more) propositions, and does not come in degrees. On the other hand there's *absolute confidence*; this includes those states that we might attribute using, for example, 'is confident to degree *x* that P', or 'is very doubtful that Q'. Absolute confidence is always directed towards a single proposition, and comes with some (possibly imprecise) degree that's often represented using values between 0 and 1.

Comparative and absolute confidence are obviously closely related to one another—the interesting question is how. According to *comparativism*, comparative confidence is strictly more fundamental than absolute confidence. Indeed, comparativists typically think that comparative confidence ought to be treated as one of the fundamental theoretical concepts in decision theory and epistemology. On this picture, absolute confidence is usually seen as a kind of 'theoretical construct', a numerical index the primary function of which is to represent where in the overall system of an agent's comparative confidences each proposition happens to sit relative to the others.

In a pair of recent papers in this journal, Stefánsson [2017; 2018] has defended comparativism, in particular against objections raised by Meacham and Weisberg [2011]. One of those objections concerns whether comparativism is capable of explaining interpersonal comparisons of confidence—for instance, whether it has the resources to make sense of one agent's having more confidence regarding some proposition P than another agent does regarding Q.

I will take it for granted that these kinds of interpersonal confidence comparisons are both meaningful and theoretically valuable, and after going over some background on comparativism and Stefánsson's proposed explanation of interpersonal comparisons in the next section, I will argue for three main conclusions. I'll first argue that Stefánsson's proposal is not compelling; indeed, it provides no reasons to think that comparativism can handle interpersonal comparisons after all (section 3). Then, I will argue that we have good reasons to think that comparativism cannot plausibly handle interpersonal comparisons (sections 4–5). And finally, I will argue that the best explanation for the interpersonal comparability of confidence involves thinking about confidence in a fundamentally different way than that which comparativists propose (section 6). We ought to see absolute confidence as a

dimensionless quantity, one which is measured not by reference to an underlying comparative confidences but via its relationship with utilities.

2. Background

For each agent α, read '$P \succcurlyeq_\alpha Q$' as saying that α's confidence regarding P is at least as great as her confidence regarding Q. We will refer to $\succcurlyeq_\alpha$ as α's *confidence ranking*. We let '$\sim_\alpha$' designate the *as much confidence* relation, and '$\succ_\alpha$' the *more confidence* relation. For the sake of simplicity, we will assume that $\succcurlyeq_\alpha$ is transitive and complete over all propositions; hence we'll treat $\sim_\alpha$ and $\succ_\alpha$ as the symmetric and asymmetric parts of $\succcurlyeq_\alpha$ respectively. Also for simplicity, I'll pretend throughout that there are only finitely many propositions. We use '$\top$' and '$\bot$' to designate a tautology and a contradiction respectively.

Say that a real-valued function on propositions f is an *order-preserving measure* of $\succcurlyeq_\alpha$ just in case

$$f(P) \geq f(Q) \text{ iff } P \succcurlyeq_\alpha Q$$

That is, f orders propositions numerically in the same way that $\succcurlyeq_\alpha$ orders those propositions by confidence. Next, define a *probability function*, p, as any real-valued function on propositions satisfying:

> NORMALISATION. $p(\top) = 1$
> NON-NEGATIVITY. $p(P) \geq 0$
> ADDITIVITY. If P,Q are mutually exclusive, then $p(P \vee Q) = p(P) + p(Q)$

Say that $\succcurlyeq_\alpha$ is *coherent* just in case at least one probability function is an order-preserving measure of $\succcurlyeq_\alpha$; and furthermore say that $\succcurlyeq_\alpha$ is *continuous* just in case no more than one is probability function is an order-preserving measure of $\succcurlyeq_\alpha$.

It has long been known that if $\succcurlyeq_\alpha$ is coherent and continuous, then it is in principle possible for comparativists to give some potential meaning to the idea of *distances* in degrees of confidence. The key observation relates primarily to ADDITIVITY, which implies that if $\succcurlyeq_\alpha$ is coherent, then the *disjunction of mutually exclusive propositions* can be treated as a kind of qualitative analogue of addition with respect to $\succcurlyeq_\alpha$. Given this, comparativist can (and usually do) follow a standard methodology from the theory of measurement to provide truth conditions for claims about ratios of differences between degrees of confidence entirely in terms of comparative confidences.[1]

To see how this would go in practice, assume that $\succcurlyeq_\alpha$ is coherent and continuous. Then we can say, for instance,

> TWICE DISTANCE. Where $P \succcurlyeq_\alpha Q$ and $R \succcurlyeq_\alpha S$, the distance between α's confidence in P and Q is *at least twice* the distance between R and S, if there are X,Y,Z such that:
> 1. $X \succcurlyeq_\alpha (Y \vee Z)$ and $Y \sim_\alpha Z$
> 2. $P \sim_\alpha (Q \vee X)$ and $R \sim_\alpha (S \vee Y)$
> 3. Y,Z are mutually exclusive, as are Q,X, and S,Y

[1] For detailed discussions of this methodology aimed at philosophical audiences, see [Fine 1973: 68ff], [Stefánsson 2017; 2018], and [Elliott 2020a; 2020b]; for a formal treatment, see [Krantz et al. 1971: 199–21]. The same methodology can also be used to give truth conditions for claims about ratios (not just ratios of differences) whenever $\succcurlyeq_\alpha$ is coherent and continuous. This is not noted in [Stefánsson 2017], but the fact is exploited in [Stefánsson 2018]. It won't make any difference to my arguments whether we think that confidence is measurable on nothing stronger than an interval scale, or if we think that it's measurable on a ratio scale. I focus on ratios of differences only because that's the focus in [Stefánsson 2017].

To flesh that out: where these conditions are satisfied, then the comparativist would typically say that the confidence α has in X is just the amount of confidence one would need to 'add' to her confidence in Q to get her confidence in P. Since α has the same confidence regarding the disjoint Y and Z, and at least as much confidence in X as in YVZ, the comparativist will say that α has at least twice as much confidence in X as in Y. Given that, and since the confidence she has in Y is just the distance between her confidence in R and in S, the result is that the distance between P and Q is at least twice the distance between R and S.

Now it's crucial to note here that TWICE CONFIDENCE makes no mention of how $\succcurlyeq_\alpha$ is measured. If f is an order-preserving measure of $\succcurlyeq_\alpha$, then it must accurately represent that the conditions stated in TWICE CONFIDENCE are satisfied. But an order-preserving measure f need not be such that the difference between the values assigned to P and to Q is *at least twice* the difference between the values assigned to R and to S. We should like a measure that does this; hence, say that f is an *interval-preserving measure* of $\succcurlyeq_\alpha$ whenever it is an order-preserving measure of $\succcurlyeq_\alpha$ and *also* adequately represents what we've determined to be the truth conditions for claims about ratios of differences in the desired form.

For example, where $\succcurlyeq_\alpha$ is coherent, and the probability function p is an order-preserving measure of it, then p will be an interval-preserving measure of $\succcurlyeq_\alpha$. From conditions 1,3 of TWICE CONFIDENCE,

$p(\mathrm{X}) \geq p(\mathrm{Y}) + p(\mathrm{Z})$, and $p(\mathrm{Y}) = p(\mathrm{Z})$
$\therefore p(\mathrm{X}) \geq 2p(\mathrm{Y})$

And then from 2, 3,

$p(\mathrm{P}) = p(\mathrm{Q}) + p(\mathrm{X})$, and $p(\mathrm{R}) = p(\mathrm{S}) + p(\mathrm{Y})$
$\therefore p(\mathrm{P})\text{-}p(\mathrm{Q}) \geq 2[p(\mathrm{R})\text{-}p(\mathrm{S})]$

If $\succcurlyeq_\alpha$ is also continuous, then p will be the unique interval-preserving measure of $\succcurlyeq_\alpha$ *on the 0-to-1 interval* (that is, the extremities of $\succcurlyeq_\alpha$ will be assigned 0 and 1, with all other values falling between); furthermore, f will be an interval-preserving measure of $\succcurlyeq_\alpha$ if and only if f is some positive affine transformation of p.[2]

On the basis of these facts, Stefánsson [2017] argues for the following:

(a) The thesis of *probabilism* amounts to the claim that ideally rational agents will have coherent confidence rankings.
(b) Comparativism can explain *distances* in degrees of confidence, at least for agents whose confidence rankings are coherent and continuous.
(c) Comparativism can explain *interpersonal confidence comparisons*, at least between agents whose confidence rankings are coherent and continuous.

I have discussed (a) and (b) elsewhere, and I think there are good reasons to doubt both [Elliott 2020_b]. But I'm not going to discuss either of them directly in this paper, so let's assume for now that they're both true. How do we get from there to interpersonal comparisons?

[2] That is, g is a positive affine transformation of f just in case for all P, $g(\mathrm{P}) = f(\mathrm{P})\mathrm{r} + \mathrm{c}$, for $\mathrm{r} > 0$ and any constant c. Except in the special case where $\mathrm{c} = 0$, any positive affine transformation of a probability function p will violate ADDITIVITY. Almost all positive affine transformations of a probability function will therefore not preserve ratios between the values that function assigns—nevertheless, they will preserve ratios of differences, and that's all we need.

Well, Stefánsson writes the following (with notation altered for consistency):

> … let me explain why we Comparativists need not give up interpersonal facts about strength of belief, contrary to what Meacham and Weisberg claim. That is, we can make sense of claims like 'α is more confident that it will rain than β is' in terms of α's and β's comparative belief relations. [2017: 581]

The proposed explanation proceeds as follows.

> It is generally assumed that … subjective probabilities (which represent strengths of belief) are interpersonally comparable … The crucial difference between desires and beliefs in this regard is the widely held assumption that any two rational people believe equally strongly whatever they fully believe (such as a tautology), and, similarly, believe equally strongly whatever they believe least of all… [2017: 581]

In this passage, for α to *fully believe* that P just means that P is maximal in $\succcurlyeq_\alpha$. So the 'widely held assumption' is that if α and β are rational—specifically, in the sense of having coherent confidence rankings $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$—then if P sits at the top (bottom) of $\succcurlyeq_\alpha$ and Q sits at the top (bottom) of $\succcurlyeq_\beta$, then α's confidence regarding P is not only comparable with but equal to β's confidence regarding Q. Let's refer to this as assumption as MIN-MAX EQUALITY. Stefánsson offers no argumentative support for MIN-MAX EQUALITY, and if you're worried about whether *comparativists* can take it for granted in the present dialectical context, then good: you should be. But we'll come back to that soon enough.

So now suppose that $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$ are not only coherent but also continuous, with p_α and p_β being the probability functions that represent $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$ respectively. Accordingly, for both agents and for any P, there is a well-defined notion of distance between P and $\top$, the latter of which will always sit at the very top of both α's and β's confidence rankings. Thus, Stefánsson notes that if MIN-MAX EQUALITY is true,

> … we might compare the degree to which α believes P with the degree to which β believes Q, by comparing the distance between P and the tautology according to α with the distance between Q and the tautology according to β. [2017: 582]

The suppressed premise here is that since α and β have the same degree of confidence as one another for $\top$ and for $\bot$, the *distance* between $\top$ and $\bot$ will be the same for each—and therefore any fraction of that distance will likewise be equal. Hence, α's confidence in P is at least as great as β's confidence in Q just in case

$$p_\alpha(\top)\text{-}p_\alpha(\mathrm{P}) \leq p_\beta(\top)\text{-}p_\beta(\mathrm{Q}),$$

which is exactly whenever $p_\alpha(\mathrm{P}) \geq p_\beta(\mathrm{Q})$. Furthermore,

> The result of the above comparison is the same across different numerical models of α's and β's comparative beliefs. That is, if α believes P more strongly than β believes Q according to one of these models, then the same holds according to all of these models. […] And (to repeat) it is a good general principle to accept as real any feature that is shared by all models of a real phenomenon. Hence, since all models of rational

> comparative belief relations agree when it comes to interpersonal comparisons, I suggest that we Comparativists take such features to be real… [2017: 582]

Note the implication here: p_α and p_β belong to 'the same numerical model', and because the same kind of comparison can consistently be made across 'all models', they therefore count as 'real'. Stefánsson doesn't explain what he means in describing two functions as belonging to the same model, but the idea seems to be this:

> SAME MODEL. Where $\succcurlyeq_\alpha$, $\succcurlyeq_\beta$ are coherent and continuous, f and g belong to the same model iff, relative to the same n-to-m interval, f and g are the unique interval-preserving measures of $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$ respectively

Thus, p_α and p_β belong to the same model. If we were to apply some positive affine transformation to, say, p_α but not p_β, then we'd end up with *different* models for α and β, which would invalidate drawing any interpersonal comparisons between them on the basis of those models. (Compare [List 2003: 232–4] on interpersonal level and unit comparisons.)

For example, for any real value r, let $t(r) = 9r + 1$. Where previously we might have said that α has less confidence in P than β does in ⊤ because

$$p_\alpha(\mathrm{P}) = 0.5 \;<\; p_\beta(\top) = 1,$$

if we apply the transformation t to p_α but not to p_β, then

$$t[p_\alpha(\mathrm{P})] = 5.5 \;>\; p_\beta(\mathrm{Q}) = 1$$

To *re*-validate the comparisons, we just need to apply t to both p_α and p_β at once, which will put the resulting measures on the same 1-to-10 interval:

$$t[\mathrm{p}_\alpha(\mathrm{P})] = 5.5 \;<\; t[p_\beta(\mathrm{Q})] = 10$$

Hence, again: f and g belong to 'the same model' just in case they're the unique interval-preserving measures of $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$ on the same n-to-m interval—and as above, 'any feature that is shared by all models of a real phenomenon' is itself 'real'.

3. Comparing Mass and Volume

Almost all of the heavy-lifting in the foregoing proposal is being done by MIN-MAX EQUALITY, and inasmuch as that assumption is left unjustified then it cannot rightly be called an explanation of why comparativists need not give up interpersonal comparisons of confidence. Perhaps you might choose to call the proposal an *incomplete* or *partial* explanation—but if that's what it is, then the part we've been given is not the part we should be worried about.

To help make this clearer, consider a parody explanation of mass-volume comparisons.[3] Imagine a finite Newtonian universe, Δ, that consists fundamentally of some array of non-pointlike atoms. The non-atomic objects of this universe are the arbitrary mereological sums of atoms. There are two special objects worth highlighting: the 'null' object ∅, or the empty arrangement of atoms; and the universal sum, Δ itself. Let '$\succcurlyeq_{\mathbf{m}}$' and '$\succcurlyeq_{\mathbf{v}}$' denote the *is at least as massive as* and *is at least as voluminous as* relations respectively. Obviously, ∅ will sit at

[3] This is an example I will come back to again later in the paper.

the bottom of both $\succcurlyeq_{\mathbf{m}}$ and $\succcurlyeq_{\mathbf{v}}$, while Δ will sit at the top. Other than that, $\succcurlyeq_{\mathbf{m}}$ and $\succcurlyeq_{\mathbf{v}}$ are two distinct orderings corresponding to two very different physical quantities.

We can define a single 'addition' operation which operates same way for mass and volume: if two objects o_1 and o_2 share no parts ($o_1 \sqcap o_2 = \emptyset$), then the mass of their mereological sum ($o_1 \sqcup o_2$) will be the sum of their individual masses, just as the volume of their mereological sum will be the sum of their volumes. We can thus define two functions $f_{\mathbf{m}}$ and $f_{\mathbf{v}}$ which measure $\succcurlyeq_{\mathbf{m}}$ and $\succcurlyeq_{\mathbf{v}}$ respectively, which are such that for all objects o_1,o_2,

($\mathrm{i}_{\mathbf{m}}$) $\quad o_1 \succcurlyeq_{\mathbf{m}} o_2$ iff $f_{\mathbf{m}}(o_1) \geq f_{\mathbf{m}}(o_2)$
($\mathrm{ii}_{\mathbf{m}}$) $\quad f_{\mathbf{m}}(\Delta) = 1$ and $f_{\mathbf{m}}(o_1) \geq 0$
($\mathrm{iii}_{\mathbf{m}}$) $\quad$ if $o_1 \sqcap o_2 = \emptyset$, then $f_{\mathbf{m}}(o_1 \sqcup o_2) = f_{\mathbf{m}}(o_1) + f_{\mathbf{m}}(o_2)$

And

($\mathrm{i}_{\mathbf{v}}$) $\quad o_1 \succcurlyeq_{\mathbf{v}} o_2$ iff $f_{\mathbf{v}}(o_1) \geq f_{\mathbf{v}}(o_2)$
($\mathrm{ii}_{\mathbf{v}}$) $\quad f_{\mathbf{v}}(\Delta) = 1$ and $f_{\mathbf{v}}(o_1) \geq 0$
($\mathrm{iii}_{\mathbf{v}}$) $\quad$ if $o_1 \sqcap o_2 = \emptyset$, then $f_{\mathbf{v}}(o_1 \sqcup o_2) = f_{\mathbf{v}}(o_1) + f_{\mathbf{v}}(o_2)$

We can thus construct a notion of *distance* for both mass and volume using the very same methodology comparativists propose for defining distances in confidence [Krantz et al. 1971]. Furthermore, $f_{\mathbf{m}}$ and $f_{\mathbf{v}}$ belong to the 'same model', as each is the unique interval-preserving measure of their respective orderings on the same 0-to-1 interval.

Of course, none of this gives us any reason at all to think that mass and volume are *comparable*. But suppose I now want to explain how we can in fact make mass–volume comparisons, and to get the ball rolling I'm going to help myself to a little assumption:

> MASS-VOLUME EQUALITY. Δ has as much mass as it does volume, and $\emptyset$ has as much mass as it does volume

Since Δ's mass just *is* its volume, and $\emptyset$'s mass *is* its volume, the distance between Δ's mass and $\emptyset$'s mass is the distance between Δ's volume and $\emptyset$'s volume—so any fraction of that distance will be equal. Thus, if

$$f_{\mathbf{m}}(\Delta) - f_{\mathbf{m}}(o_1) \leq f_{\mathbf{v}}(\Delta) - f_{\mathbf{v}}(o_2),$$

then we say that o_1's mass is at least as great as o_2's volume, which will be whenever $f_{\mathbf{m}}(o_1) \geq f_{\mathbf{v}}(o_2)$; and all such comparisons will be preserved whenever mass and volume are measured on 'the same model'.

So *if* MASS-VOLUME EQUALITY is true, *then* we can make sense of mass-volume comparisons. But that's not very interesting, and it doesn't help to support the sensibility of mass-volume comparisons in the slightest. Similarly, if MIN-MAX EQUALITY is true, then interpersonal confidence comparisons might be meaningful under certain conditions. But that conditional isn't playing any interesting role in the explanation of how interpersonal comparability might make sense in the first place. The hard part isn't to establish the conditional; it's to establish the antecedent!

What comparativists need is a justification for MIN-MAX EQUALITY (or any other posited equalities between locations in α's and β's confidence rankings). Such a justification needs to explain what's *different* between interpersonal confidence comparisons and mass-volume comparisons, and it needs to not undermine the support for comparativism more generally. Without this, the explanation laid out in section 2 is no more compelling than the parody. The

question for the remainder of this paper is therefore whether we can expect that some such justification will be forthcoming.

4. The Functional Role of Absolute Confidence

Now you might be thinking that there's an obvious difference between interpersonal confidence comparisons and mass-volume comparisons. On the one hand, it's not useful in any sense to say that $f_{\mathbf{m}}$ and $f_{\mathbf{v}}$ belong to 'the same model' precisely because mass and volume are very different physical phenomena. On the other hand, however, you might think that it's sensible to say p_α and p_β belong to the 'same model', and hence to compare between them, because they're models of similar psychological phenomena.

Well, that alone not going to be quite enough to justify MIN-MAX EQUALITY. On the comparativist's picture, p_α and p_β are interval-preserving measures of two distinct psychological quantities: there's *confidence-for-α* (underwritten by $\succcurlyeq_\alpha$), and there's *confidence-for-β* (underwritten by $\succcurlyeq_\beta$). Now $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$ are clearly similar to one another in many respects—but then so too are $\succcurlyeq_{\mathbf{m}}$ and $\succcurlyeq_{\mathbf{v}}$, so pointing out similarities in the underlying rankings won't justify MIN-MAX EQUALITY. Nor can the fact that confidence-for-α and confidence-for-β play psychologically similar roles be enough to justify that assumption. After all, it's also true that *utility-for-α* (underwritten by α's preferences) and *utility-for-β* (underwritten by β's preferences) play psychologically similar roles, and yet we certainly shouldn't take *that* as sufficient evidence that interpersonal utility comparisons are therefore meaningful.[4] So mere psychological similarity isn't going to suffice to justify MIN-MAX EQUALITY.

But maybe there's a little more that can be said in support of MIN-MAX EQUALITY on this front. The rough idea would be this: the psychological state that α is in when she has P sitting at the top of her confidence ranking plays the same functional role to the psychological state that β is in when β has P sitting at the top of her confidence ranking; and likewise for the states α and β are in when they have P sitting at the bottom of their confidence rankings. If so, then the close similarity of the functional roles might entitle us to say that they are the same psychological state—or, at the very least, that the maxima and the minima of α's and β's confidence rankings are not only comparable but indeed equal in strength. Let's call this the *same-role response*.

I do not think that the same-role response is successful. The problem, as I see it, is that the response supports MIN-MAX EQUALITY only at the cost of undermining comparativism more generally. To get clearer on this, let's flesh the idea out in a bit more detail. On the usual comparativist's picture, if p_α and p_β are the unique interval-preserving measures of α's and β's confidence rankings on the 0-to-1 scale, then these functions can be plugged into our standard numerical models of decision-making to predict α's and β's utilities for certain kinds of gambles relative to their utilities for the outcomes of those gambles. Taking a simplified version of ordinary expected utility theory for our main example, the utility u_α that a rational agent like α assigns to a gamble,

$$\Gamma = \langle \text{Q if P, R otherwise} \rangle,$$

[4] To be explicit: I am assuming that *utility-for-α* is an interval-preserving measure of α's preferences; likewise for β. I'm therefore assuming that rational agents have preferences that are measurable as such. This should be generally uncontroversial. I am *not* arguing that interpersonal utility comparisons are meaningless. I do happen to think that interpersonal utility comparisons are meaningless, but right now I'm only noting that the mere fact that *utility-for-α* and *utility-for-β* are psychologically similar does not imply that they're comparable.

is a function of the utilities she assigns to Q and R, and the degree of confidence she assigns to P:[5]

$$u_\alpha(\Gamma) = p_\alpha(\mathrm{P})u_\alpha(\mathrm{Q}) + [1\text{-}p_\alpha(\mathrm{P})]u_\alpha(\mathrm{R})$$

Where α prefers Q to R, we can rearrange this to give:

$$u_\alpha(\Gamma) = u_\alpha(\mathrm{R}) + p_\alpha(\mathrm{P})[u_\alpha(\mathrm{Q})\text{-}u_\alpha(\mathrm{R})]$$

This is just another way to say that the utility α assigns to Γ will sit $p_\alpha(\mathrm{P})$ of the distance from α's utility for R to her utility for Q. So, if P is maximal in $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$ then $p_\alpha(\mathrm{P}) = p_\beta(\mathrm{P}) = 1$, and α and β both will be indifferent between Q and Γ. Likewise, if P is minimal in $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$ then $p_\alpha(\mathrm{P}) = p_\beta(\mathrm{P}) = 0$, and they will be indifferent between R and Γ. Supposing that all this is correct, then I think it would be perfectly plausible to say that a proposition's sitting at the top (bottom) of α's confidence ranking plays the same functional role in relation to utilities for α as a proposition's sitting at the top (bottom) of β's confidence ranking plays for β.

But you know what other states would also play the same functional roles across α and β? If P sits *half* of the way between ⊤ and ⊥ on α's confidence ranking, then the utility α assigns to Γ will sit half of the distance from $u_\alpha(\mathrm{R})$ to $u_\alpha(\mathrm{Q})$. Similarly, if P sits a *quarter* way between ⊤ and ⊥ on α's confidence ranking, then the utility α assigns to Γ will sit a quarter of the distance from $u_\alpha(\mathrm{R})$ to $u_\alpha(\mathrm{Q})$.

In general, the same-role response has us identify *some* of α's and β's psychological states by virtue of those states' functional roles: being maximally-ranked in $\succcurlyeq_\alpha$ and $\succcurlyeq_\beta$ counts equally as 100% *confidence* for both α and β because (according to the usual comparativist theory) those states behave in the same way with respect to utilities; and likewise for 0% *confidence*, *mutatis mutandis*. But why did we stop there? What we've been describing are specific instances of much more general way of defining any state of absolute confidence directly by its functional relationship with utilities. That is, α and β are both *x% confident that* P just when they're in a state the functional role of which leads them to assign a utility to Γ that is *x%* of the distance from the utility they assign to R to the utility they assign to Q. Or in other words, where α still prefers Q to R,

$$p_\alpha(\mathrm{P}) = [u_\alpha(\Gamma)\text{-}u_\alpha(\mathrm{R})]/[u_\alpha(\mathrm{Q})\text{-}u_\alpha(\mathrm{R})]$$

But now comparative confidence has dropped out of the picture. To characterise α's degree of confidence regarding P in *this* way I don't need to know where P sits relative to other propositions in $\succcurlyeq_\alpha$. And that's because once we start characterising confidence by reference to these kinds of functional roles, α's confidence regarding P is not in the first instance being treated as an index that represents the location of P relative to other propositions in a confidence ranking, but rather as a measure, roughly, of the degree to which α is willing to bet on P.[6]

[5] I'm here presupposing in what follows that there are some Q and R such that α prefers Q to R; and to keep things simple I'm assuming that α is indifferent between Q and (Q∧P), and between R and (R∧¬ P).

[6] Following Eriksson and Hájek [2007], you might worry here about so-called 'Zen monk' cases, or agents who are indifferent amongst all things. I have responded to this problem elsewhere [Elliott 2019]. In short: a functional characterisation of α's confidence states isn't given in terms how those states interact with her actual utilities\preferences, but their potential interactions with different utility\preference states she could be in. If α is actually indifferent amongst all things, then she can still be in a state the typical causal role of which would only become apparent if she were to no longer be universally indifferent.

(I'll note, by the way, that the point here doesn't rest on the simplified decision theory I've used for the example. Compare the simple expected utility theory I've been presupposing with something like Buchak's [2013] risk-weighted utility theory. Unlike the expected utility model, which says:

$$u_\alpha(\Gamma) = u_\alpha(\mathrm{R}) + p_\alpha(\mathrm{P})[u_\alpha(\mathrm{Q})\text{-}u_\alpha(\mathrm{R})]$$

Buchak posits a (strictly increasing) risk function, r_α: [0,1] ↦ [0,1], with $r_\alpha(0) = 0$ and $r_\alpha(1) = 1$, that is intended to represent α's attitude towards risk; and then asks us to calculate utilities like so:

$$u_\alpha(\Gamma) = u_\alpha(\mathrm{R}) + r_\alpha[p_\alpha(\mathrm{P})][u_\alpha(\mathrm{Q})\text{-}u_\alpha(\mathrm{R})]$$

Assuming p_α, u_α, and r_α are understood to represent distinct psychological phenomena with distinct functional roles, we can characterise a given state of absolute confidence by its functional relationships with utilities *and* risk attitudes:

$$p_\alpha(\mathrm{P}) = r_\alpha^{-1}([u_\alpha(\Gamma)\text{-}u_\alpha(\mathrm{R})]/[u_\alpha(\mathrm{Q})\text{-}u_\alpha(\mathrm{R})])$$

If α is risk-neutral, then $r_\alpha^{-1}(\mathrm{n}) = \mathrm{n}$, and there's no difference between a gamble's risk-weighted utility and its expected utility. So given risk-weighted utility theory as our underlying decision model we could treat '$p_\alpha(\mathrm{P}) = x$' as a measure of α's willingness to bet on P *if she were risk neutral*. But the key point is just that *what it is* for α be confident that P to degree x can be characterised in terms of that state's functional relationships with other states posited within some decision theory, including at least but perhaps not limited to its relationship with utilities, and without referring to P's relative location in a confidence ranking.)

The argument here is that we don't get to pick and choose when we appeal to similar functional roles: if we're going to use the functional roles of confidence states in connection with other psychological phenomena to characterise what it is for agents to have 100% confidence that P, or 0% confidence that P, then we should recognise when the same functionalist definitions can be used to characterise what it is for those agents to have x% confidence that P for any x between 0 and 100—that is, *without* at any point mentioning the agent's comparative confidences—and not appeal to functionalism only when it suits the theory we're trying to support. If the same-role response does anything to support MIN-MAX EQUALITY, then it's only at the cost of undermining comparativism more generally.[7]

5. When Comparisons are Meaningful

So perhaps it's not so easy to justify MIN-MAX EQUALITY. And in fact I think we have good general reasons to think that no compelling justification will be forthcoming. The key question to ask is: when does it make sense to draw comparisons between quantities?

[7] I'll flag here that I think there are further problems with the same-role response. I've been granting for the sake of argument that if two agents have identical coherent and continuous confidence rankings, then they have identical absolute confidences. But *that's* a commitment of Stefánsson's comparativism, not a self-evident truth. It is, at least arguably, conceptually possible for two agents to have identical confidence rankings and yet attach different absolute confidences to propositions at the same 'locations' within their respective rankings (including the minima and maxima). This includes cases where the differences in absolute confidence between the agents are systematically reflected by differences in their preferences as predicted by an underlying decision theory, and are thus functionally distinct according to that theory. This is, however, a more general problem for comparativism that I've discussed elsewhere [Elliott 2020b], I don't want to dwell on it further here.

Well, it's precisely when those comparisons aren't reliant on any unforced or arbitrary choices relating to the format of the representations used. And across all clearly meaningful instances of comparability—both with respect to physical quantities as well as biological, psychological, or sociopolitical quantities—there are four general kinds of case where this is true. So the goal of this section is to argue that if comparativism were true and interpersonal confidence comparisons were indeed meaningful, then they would be quite unlike any of these four standard kinds of cases.

(That's consistent with interpersonal confidence comparisons being a *unique case*, of course—but then wouldn't it be so much nicer to have a theory on which interpersonal confidence comparisons *aren't* fundamentally distinct from other forms of quantitative comparisons we find in the sciences?)

Let $\mathbf{q}_1$ and $\mathbf{q}_2$ designate two not necessarily distinct quantities for which it makes some sense to talk about 'distances' and ratios thereof. Every such quantity $\mathbf{q}$ induces an ordering, $\succcurlyeq_{\mathbf{q}}$, over the kinds of things for which that quantity is attributable; for example, mass and volume induce the orderings $\succcurlyeq_{\mathbf{m}}$ and $\succcurlyeq_{\mathbf{v}}$ over the space of concrete objects respectively. So let $f_{\mathbf{q}1}$ be an interval-preserving measure of $\succcurlyeq_{\mathbf{q}1}$, and likewise let $f_{\mathbf{q}2}$ be an interval-preserving measure of $\succcurlyeq_{\mathbf{q}2}$. Then the four kinds of circumstances where it unambiguously makes sense to draw $\mathbf{q}_1$–$\mathbf{q}_2$ comparisons from the values assigned by $f_{\mathbf{q}1}$ and $f_{\mathbf{q}2}$ are:

C1. $\mathbf{q}_1 = \mathbf{q}_2$ and $f_{\mathbf{q}1} = f_{\mathbf{q}2}$
C2. $\mathbf{q}_1 = \mathbf{q}_2$ and $f_{\mathbf{q}1} \neq f_{\mathbf{q}2}$, but we know how to translate between $f_{\mathbf{q}1}$ and $f_{\mathbf{q}2}$
C3. $\mathbf{q}_1 \neq \mathbf{q}_2$, but both the $\mathbf{q}_1$-facts and the $\mathbf{q}_2$-facts can be re-expressed in terms of a single theoretically more basic quantity, $\mathbf{q}_3$
C4. $\mathbf{q}_1 \neq \mathbf{q}_2$, but both $\mathbf{q}_1$ and $\mathbf{q}_2$ are dimensionless

In the remainder of this section I'll describe these in turn, and I'll argue that if comparativism were true then interpersonal confidence comparisons could fit none of these patterns.

5.1 Cases C1 and C2

These are the simplest and most obvious cases. An example of C1 would be if we have a single quantity, *mass*, measured on a single numerical scale, *kilograms*; and in this case we can draw *mass-mass* comparisons between any two objects by reading the comparisons directly off of the numerical values they're assigned on the kilogram scale. An example of C2 would then be when we have the one quantity *mass* measured on two different scales, for example, *kilograms* and *pounds*. Here we can draw *mass-mass* comparisons on the basis of the values assigned by the two scales whenever we know how to translate between those scales.

Comparativism obviously cannot directly appeal to C1 and C2. As every quantity induces an ordering over the domain appropriate to that quantity, we can use those orderings to differentiate between quantities—in the sense that if $\succcurlyeq_{\mathbf{q}1} \neq \succcurlyeq_{\mathbf{q}2}$, then $\mathbf{q}_1 \neq \mathbf{q}_2$.[8] And since $\succcurlyeq_{\alpha} \neq \succcurlyeq_{\beta}$, the comparativist is committed to saying that confidence-for-α is not the same quantity as confidence-for-β. So that rules out C1 and C2.

5.2 Case C3

Here's a simple example of C3: by stipulating the directions for *up*, *forwards*, and *across*, we can order objects by *height*, *length*, or *width*, and in that way we can make sense of these as

[8] In saying this I'm taking no stand on whether quantitative facts are anything over and above relational facts. See [Dasgupta 2013] for discussion. Even if you think there's more to the facts about a quantity than its relational facts, you'll still agree that every quantity determines some relational facts that we can then use to differentiate between them.

three distinct quantities—each corresponds to a distinct relation over concrete objects. But we usually don't think of these as interestingly distinct quantities, and the reason is that all of the relevant facts about each can be re-expressed using a single more basic quantity, *spatial distance*, plus a direction. Thus it makes perfect sense to compare o_1's height to o_2's length or to o_3's width, precisely because those comparisons reduce to more fundamental spatial distance comparisons that are then clearly meaningful under either pattern C1 or C2.

C3 is the standard pattern by which theorists will attempt to render cross-quantitative comparisons meaningful (with the exception of C4, which only applies in the case of dimensionless quantities). For example, if you want to compare aesthetic to pragmatic value, for instance, or gustatory to audible to tactile pleasure, then the usual strategy is to try to reduce both to a more basic measure—'overall value', 'overall pleasure'—under which it makes sense to trade them off against one another. And where it's not clear how to reduce distinct quantities to a common underlying measure, this is usually seen as compelling evidence of incomparability.

So can the comparativist make appeal to something like C3 to explain how confidence-for-α is comparable to confidence-for-β? Well, they'd need to show that there's a more fundamental quantity of which these are just 'aspects'—something like *confidence simpliciter*. But what grounds this more fundamental quantity, and on what basis is it measured? Not on the basis of any *individual's* confidence ranking: if confidence-for-α and confidence-for-β are going to be reducible to confidence simpliciter, then any method of measuring the latter would need to be independent of the subjective properties of α's and β's confidence rankings. But the point of comparativism was to show how absolute confidence arises for each agent out of that agent's subjective confidence ranking—so who's confidence ranking is going to be the basis for plain ol' confidence?

And this, by the way, lets us see more clearly what's so problematic about the assumption of MIN-MAX EQUALITY. Compare again the comparison of mass and volume. It's nonsense to compare these two quantities, and the ultimate reason for this is that there's no common measure to which the facts about both are reducible. The assumption of MASS-VOLUME EQUALITY amounts to stipulating a common measure out of thin air: *location relative to ∅'s mass-volume and Δ's mass-volume*. But of course you need to first establish a common measure of mass and volume *before* you can justify equating any two points between a measure of mass and a measure of volume.

So the task for comparativists is to establish the existence of a common measure. This is exactly what the same-role response does, essentially by re-expressing confidence-for-α and confidence-for-β in terms of a common measure of *confidence simpliciter* characterised by their shared functional role in relation to utilities. But the same-role response undermines comparativism as a whole, and now we can see that the basic reason for this is in fact quite general: we cannot say that confidence-for-α and confidence-for-β are both reducible to a single common (and therefore non-subjective) measure without also giving up on the idea that the facts about each agent's states of absolute confidence are grounded in the particular way that agent orders propositions by relative confidence. You can have one or the other—and if you want the latter, then you cannot have C3.

5.3 Case C4

So finally, C4. A quantity **q** is dimensionless when it is defined in terms of other quantities **q′**, **q″**, …, in such a way that the units of the latter quantities 'cancel out'. For example, the *refractive index*, n, of a substance is the ratio of the speed of light c in a vacuum in unit-distance per unit-time, to the phase velocity p of light in the medium of that substance as measured in the same units:

$$n = [c(\text{unit-distance/unit-time})]/[p(\text{unit-distance/unit-time})] = c/p$$

Because the (unit-distance/unit-time)'s in the denominator and the numerator cancel each other out, the refractive index n doesn't have 'units' in the same way that measurements of distance and duration typically do. Instead, it is a simple ratio between the two real values c and p. Likewise, the *relative density*, r, of a substance is a ratio of the density s of a given substance (in unit-mass per unit-volume) to the density m of a reference material as measured in the same units:

$$r = [s(\text{unit-mass/unit-volume})]/[m(\text{unit-mass/unit-volume})] = s/m$$

Refractive indices and *relative densities* are very different quantities from one another, and there's no more basic quantity of which both are merely 'aspects' *à la* C3, it still makes perfect sense to say that the refractive index of some medium is *greater than* its relative density. As neither quantity *has* units, the comparison is independent of any arbitrary choice of units—indeed, we're essentially just saying that one ratio is bigger than another.

We can see that the comparativist cannot appeal to C4 simply by noting that confidence-for-α would not be dimensionless if comparativism were true. Again: the comparativist's view is that the facts about an agent's absolute confidences are derived from the facts about their confidence ranking, rather than defined as a ratio of values in some further psychological quantity or quantities.

Before I close this section, let me briefly consider a possible response.[9] You might be thinking that given any interval-preserving measure f of $\succcurlyeq_\alpha$, confidence-for-α can be 'redefined' as a dimensionless ratio of differences assigned by p_α. So α believes P to degree x just in case

$$x = [p_\alpha(\mathrm{P})\text{-}p_\alpha(\bot)]/[p_\alpha(\top)\text{-}p_\alpha(\bot)]$$

And if we do the same for $\succcurlyeq_\beta$ then we have two dimensionless quantities that can now be compared. But it's important to note here that a dimensionless quantity is always defined in terms of some *other* quantity/quantities. It's nonsensical to say that confidence-for-α is a dimensional quantity measured by p_α *and* that it's 'redefinable' in terms of p_α as a dimensionless quantity. A quantity cannot be both dimensional and dimensionless, and if p_α *is* a measure of confidence-for-α, then the dimensionless quantity defined from p_α *isn't*.

What's really happening here isn't a 'redefinition' of confidence-for-α, but the defining of a new dimensionless quantity: *distance from* $\top$ *on an arbitrary interval-preserving measure of* $\succcurlyeq_\alpha$. And that's a perfectly well-defined (if not especially interesting) quantity, but it's *not* confidence-for-α. This is important, because if this redefinition strategy were sensible then we could do the very same thing for the case of mass and volume—that is, just let o's 'mass' be equal to

$$[f_{\mathbf{m}}(o)\text{-}f_{\mathbf{m}}(\emptyset)]/[f_{\mathbf{m}}(\Delta)\text{-}f_{\mathbf{m}}(\emptyset)]$$

and let its 'volume' be equal to

$$[f_{\mathbf{v}}(o)\text{-}f_{\mathbf{v}}(\emptyset)]/[f_{\mathbf{v}}(\Delta)\text{-}f_{\mathbf{v}}(\emptyset)],$$

[9] I thank an anonymous referee for raising this suggestion.

and voilà we can now compare the dimensionless 'mass' and 'volume'! But of course we cannot use this strategy to make sense of mass–volume comparisons, because 'mass' and 'volume' aren't mass and volume.

To reiterate what I said above, any strategy for making sense of interpersonal confidence comparisons ought to show what's *different* between them and mass-volume comparisons; otherwise, applicability to the latter stands as a *reductio* of the former.

6. Confidence as Dimensionless

I think the right way to understand absolute confidence is by its relationship with utilities and preferences, along the lines that I described in section 4. For lack of a better name, let's just call this the *functionalist view*. This is a view that I've defended in several works [Elliott 2017, 2019, 2020a]. To close the paper, let me say a few things specifically in relation to how the view handles interpersonal comparisons of confidence.

First, on the functionalist's picture, α's absolute confidence for any proposition P is a *dimensionless* quantity. The value $p_\alpha(P)$ is a ratio of two distances in utility—the distance between $u_\alpha(\Gamma)$ and $u_\alpha(R)$, and the distance between $u_\alpha(Q)$ and $u_\alpha(R)$:

$$p_\alpha(P) = [u_\alpha(\Gamma)-u_\alpha(R)]/[u_\alpha(Q)-u_\alpha(R)]$$

Since the denominator and the numerator have the same units (*utils*), they'll cancel each other out, leaving us with the dimensionless $p_\alpha(P)$. And since p_α and p_β measure dimensionless quantities for both α and β, it makes perfect sense to compare across them. Note that this is true regardless of what specific scales we use to measure α's and β's utilities, so long as that measure is interval-preserving; and consequently, at no point did we need to assume that α's and β's utilities are interpersonally comparable. Moreover, since p_α and p_β are defined in the same way by reference to their similar functional roles in relation to α's and β's utilities respectively, it's not only meaningful but also *useful* to compare across them. For instance, if $p_\alpha(P) > p_\beta(Q)$, then α will be more willing to bet on P than β is on Q (ceteris paribus). This is the key insight of the same-role response, and it applies with all the more force for the functionalist view.

Interestingly, the language with which we attribute degrees of confidence also fits the pattern of dimensionless quantity attributions. To avoid ambiguity, attributions of dimensional quantities like length, mass, and temperature require specification of a unit. For instance, in most contexts we need to say '*o* has a length of 10 *meters*' or '*o* weighs 10 *kilograms*'. But because dimensionless quantities have no units we say, for example, 'water has a refractive index of 1.33', or 'wood has a relative permeability of 0.9'. Likewise, we say 'α believes *p* to degree *x*'—not 'α believes *p* with *x credals*', as one ought to expect if confidence were a dimensional quantity as the comparativist proposes. Or more instructively: we naturally understand and describe confidence in terms of *percentages*, which are just another way of representing dimensionless ratios.

So, the functionalist view has a neat explanation of interpersonal confidence comparisons. The explanation does not rely on any arbitrary choice of units, nor on any controversial presuppositions of interpersonal comparability, or on questionable equivalences between the relative positions of propositions on an agent's confidence ranking and how confident the agent is regarding those propositions. It also fits nicely with the ways we talk about confidence, both in our formal theories and in everyday speech. Compared to comparativism, then, the functionalist view has a lot going for it when it comes to explaining interpersonal comparisons. But even if you don't like my proposed alternative, it's clear enough that comparativists are in need of a better response to the problem of interpersonal confidence comparisons.

Acknowledgements: Thanks are due to Nick DiBella for helpful discussions on the topic, and to anonymous referees.

REFERENCES

Buchak, L. 2013. *Risk and Rationality*. New York: Oxford University Press.

Dasgupta, S. 2013. Absolutism vs Comparativism About Quantity, in *Oxford Studies in Metaphysics, Volume 8*, ed. Karen Bennet and Dean W. Zimmerman, New York: Oxford University Press: 105–148.

Elliott, E. 2019. Betting Against the Zen Monk, *Synthese*. URL = <https://doi.org/10.1007/s11229-019-02308-4>

Elliott, E. 2020a. 'Ramseyfying' Probabilistic Comparativism, *Philosophy of Science* 87/4, 727–754.

Elliott, E. 2020b. Comparativism and the Measurement of Partial Belief, *Erkenntnis*. URL = <https://doi.org/10.1007/s10670-020-00329-x>

Eriksson, L. and A. Hájek 2007. What are degrees of belief?, *Studia Logica* 86/2, 183–213.

Fine, T. L. 1973. *Theories of Probability: An Examination of Foundations*. New York: Academic Press.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky 1971. *Foundations of measurement, Volume I: Additive and polynomial representations*. New York: Academic Press.

List, C. 2003. Are Interpersonal Comparisons of Utility Indeterminate?, *Erkenntnis* 58/2, 229–260.

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

Stefánsson, H. O. 2017. What Is 'Real' in Probabilism?, *Australasian Journal of Philosophy* 95/3, 573–587.

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

Institutional Affiliation:
School of Philosophy, Religion and History of Science
University of Leeds