Does "better than" have a uniform scale?

Erich Rast

Nova Institute of Philosophy, Universidade Nova de Lisboa

ArgLab Open Seminar, March 11, 2022

Overview

Value Adjectives

1. Value adjectives like *brave* and *good* often are gradable, i.e., have a comparative and superlative form in addition to the positive.
2. This is not generally the case. See e.g. **lewder*, *?more excellent*.
3. One common construction is to derive a base form such as *good* from the comparative *better than*, which elicits an underlying value structure.
4. This is not the only option. There are also constructions for deriving *better than* from uses of *good*, and similar for other value adjectives.
5. Problem: Given the first approach based on value structure, can we assign numbers to value comparisons?

▷ **This question is generally studied in measurement theory.**

Utility Function

A utility function $u : D \to \mathbb{R}$ maps items in a domain D to real numbers. There are often additional restrictions:

1. Exclude disvalue: Restriction to only positive real numbers.
2. Normalize utilities: Restriction to real numbers in the unit interval $[0, 1]$.
3. "More is better": Sometimes only monotonically increasing functions are allowed.
4. Technical conditions: Conditions like continuity, smoothness, twice differentiable, etc.
5. Restrictions of the shape of the curve: concavity (diminishing marginal utility), convexity, linearity, etc.

▷ **I do not make these assumptions but it's worth noting that for certain applications in economics strong assumptions are sometimes needed.**

Value Relations

Value relations are representations of value comparisons between items (states of affairs, objects, propositions, etc.). They are often thought to be binary relations over the domain of items, i.e., $R \subseteq D \times D$.

1. Define weak *better than or equally good* as R, then define *better than* as $P(x, y) := R(x, y) \ \& \ \neg R(y, x)$ and *equally good* as $I(x, y) := R(x, y) \ \& \ R(y, x)$. But we can also define R from P and I.
2. In the "standard view" in economics, a value relation R is complete over D: $\forall x, y \in D : R(x, y) \vee R(y, x)$, reflexive $\forall x \in D : R(x, x)$ and transitive $\forall x, y, z \in D : \big([R(x, y) \ \& \ R(y, z)] \rightarrow R(x, z)\big)$.
3. This is highly contested:
   - Many ethicists question completeness due to moral dilemmas and other problems with comparing items.
   - Temkin, Rachels, and others have argued that *better than* comparisons might not transitive (or other common assumptions have to be given up).

▷ **Note: Actual value comparisons are always multidimensional, even for intrinsic value (Rast 2020; 2022). But for the current purpose talking about monist value representations and overall goodness suffices.**

Representation Theorems

Representation theorems relate a numerical representation with a value relation. The numerical representation must *represent* the value relation. In the simplest case, the condition is as follows:

$$u(a) \geq u(b) \Leftrightarrow R(a, b) \tag{1}$$

Or, in terms of *better than* expressed by P and *equally good* expressed by I:

$$u(a) > u(b) \Leftrightarrow P(a, b) \tag{2}$$
$$u(a) = u(b) \Leftrightarrow I(a, b) \tag{3}$$

▷ **This holds for the "standard view" in finite and countable domains; technical conditions on R are needed to ensure this holds for uncountable infinite domains. More importantly: Other value relations have more complicated representation theorems or are not representable by a single utility function at all.**

Scale Types

Common scale types were defined by psychologist Stevens (1946). Often used as a reference is Krantz et al. Vol. I-III (1971, 1989, 1990); I base my characterizations on the excellent book by Roberts (1979, 1985). Note that measurement theory is really about empirical measurement, whereas our value judgments are generally not measured but posited. There is a huge difference in perspective!

In economics, measurement of preferences became a prevalent problem in the work of Vilfredo Pareto (1900; 1906) and later in revealed preferences theory of Samuelson (1947). That is because economists stipulated utility functions during 19th Century under influence of early utilitarians, and Pareto started to question their empirical foundation.

Common Scale Types

Scale types can be defined by admissible transformations from one utility function into another. If $u(x)$ represents R, then a function $u'(x)$ will also represent R and is on the same scale if it can be obtained by one of the admissible transformations (Roberts 1985: 66-7):

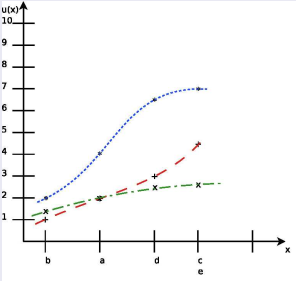
- **Ordinal Scale**: $u'(x) = f(u(x))$ for any increasing function f
  Example: scale for quality from 10 "best" to 1 "worst"
- **Interval Scale**: $u'(x) = au(x) + b$ for $a > 0$ and any b
  Example: temperature in Celsius and Fahrenheit
- **Ratio Scale**: $u'(x) = au(x)$ for $a > 0$
  Example: mass, temperature in Kelvin
- **Nominal Scale**: $u'(x)$ is a one-to-one transformation of $u(x)$
  Example: numbers on shirts in a soccer team, passport numbers; numbers only serve as labels
- **Absolute Scale**: $u'(x) = u(x)$
  Example: counting

Example 1

$D = \{a, b, c, d, e\}$, Relation R defined by $cIePdPaPb$. Best score: 10, worst score 1. u, u', and u'' define the same ordinal scale and represent R.

Ordinal Scale

	u	u'	u''
a	2.0	4.0	$\sqrt{4}$
b	1.0	2.0	$\sqrt{2}$
c	4.5	7.0	$\sqrt{7}$
d	3.0	6.5	$\sqrt{6.5}$
e	4.5	7.0	$\sqrt{7}$



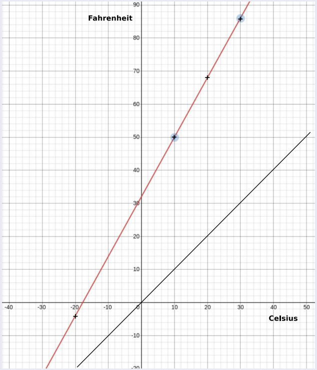
Example 2

$D = \{a, b, c, d, e\}$, Relation R defined by $cIePdPaPb$. Functions now denote temperatures.

Interval Scale

$$\begin{array}{cccccc} F(x) & = & 1.8 & C(x) & + & 32 \\ u'(x) & = & a & u(x) & + & b \end{array}$$

	Celsius	Fahrenheit
a	10	50
b	-20	-4
c	30	86
d	20	68
e	30	86



Types of Arguments for Scales

1. Linguistic Arguments
2. Technical Arguments / Arguments from Practicality
3. Normative Arguments

▷ **In the literature, the last two categories are often mixed. It can be hard to discern practical from normative considerations in this area.**

Linguistic Arguments

Lassiter (2017), Soria-Ruiz (2021):

Intensifiers like *much, a little, way, very*:

(1) *Your plan is much better than Jeff's.*

(2) *This knife is way better than that one.*

▷ **meaningful on interval and ratio scales**

'Ratio' Modifiers like *two times*, $1.38\times$, *a hundred times*:

(3) *Salad is seven times healthier than French fries.*

(4) *This proposal is a hundred times better than the previous one.*

▷ **meaningful on ratio scales**

Interval vs. Ratio Scale

- Lassiter (2017): *better than* is on an interval scale. Uses of ratio modifiers are hyperbole, basically using the modifier as an intensifier.

  (5) *You have to be (# exactly) twice as good as Concha.*

- Soria-Ruiz (2021): Uses of ratio modifiers can be meant literally, but not all such modifiers are allowed:

  (6) *You have to be better, but not twice as good as Concha.*

  (7) #*Volunteering is 1.38x as good as donating.*

▷ **Soria-Ruiz develops a theory of round ratio scales that only allow certain round modifiers (but parameterized) for value adjectives: *twice as good*, *ten times better*, etc.**

Arguments by Practicality

These are usually implicit in uses of utility functions:

- Expected utility theory requires ratio scales because $EU(x) = u(x)P(x)$.
- Value aggregation methods often require ratio scales, e.g. weighted sum $u(x) = \sum_{i+1}^{n} u_i(x_i)w_i$.
- Common preference elicitation methods require ratio scales, e.g. asking for the certainty equivalent, finding the midpoint of a utility function.
- Ratio scales allow for the ‘full power of calculus’, advanced econometrics.

▷ **Especially the expected utility thesis and the modeling of risk attitudes seem to be very strong arguments in favor of ratio scales.**

Normative Arguments

These are hard to distinguish from the previous ones:

- Qualitative combination with qualitative notions of plausibility loose too much information.
- Qualitative value aggregation by social choice methods or the Sugeno integral are also very information lossy.
- Qualitative aggregation methods fall prey to variants of Arrow's theorem. (hierarchical preferences)
- The meaning of *better than* implies a ratio scale.
- Reflective equilibrium requires cardinal utility for full commensurability.

▷ **Maybe these aren't really normative arguments? That depends a lot on the underlying conception of rationality.**

Counter-Arguments

1. Linguistic arguments are not entirely conclusive.
2. Arguments by practicality are fallacious and based on wishful thinking.
3. Normative arguments are questionable because there are also normative arguments that speak against the idea that *better than* resides on any simple scale. (conflicting intuitions? conflicting rationality principles?)

Linguistic Arguments Are Inconclusive

Rank-based accounts of 'ratio' modifiers and intensifiers are possible. Example: $cIePdPaPb$

- Define canonical utilities in terms of rank only: $u(b) = 1$, $u(a) = 2$, $u(d) = 3$, $u(e) = u(c) = 4.5$
- x *much better than* y: $u(x) - u(y) > \delta$
  "xPy and the difference between the ranks of x and y is higher than a threshold"
- x *is 2 times better than* y: $u(x) \geq 2u(y)$ where $2u(y) = u(z)$ for some $z \in D$, "the rank of x is at least 2 times the rank of y"

▷ **This approach relaxes an absolute scale (canonical utility = rank) to allow multiplication as long as the result is a valid rank. If $n \cdot u(x)$ is not valid rank, the use of the ratio modifier is meaningless or hyperbole. Dealing with ties is tricky in this approach, and it is better to use the underlying relation directly for the definitions.**

Ratio Modifier Uses Are Questionable

I also share Lassiter's concerns. Consider:

(8) ?Bugattis are three times better than Porsches.

(9) ?Basic democracy is twice as good as representative democracy.

▷ **Even though rank-based and ratio scale based definitions could be given, it seems far stretched to claim that utterances like these can be meant literally. Only when "better than" is used as a direct proxy for an extensively measurable quantity like money might such comparisons make sense.**

Arguments by Practicality Are Fallacious

- If "better than" wasn't on a ratio scale, we couldn't apply the expected utility hypothesis to obtain evaluations of outcomes based on their utility and the probability of their occurrence.
- *Therefore, "better than" must be on a ratio scale.

▷ **Arguments like these are fallacious, based on wishful thinking. Of course, that's why also nobody really makes them explicit. But it seems that ratio scales are often assumed because they offer technical advantages.**

Normative Arguments Are Dubious

- Ordinal utilities can only be combined meaningfully with corresponding qualitative notions of plausibility (or, possibility).
- Information about quantified risk *can* be obtained and contains much more information than mere qualitative plausibility.
- So we *should* use a notion of quantified value based on cardinal utility (interval or ratio scales).

▷ **Maybe these are still wishful thinking? The idea would be that we posit ratio scale utilities and elicit them from agents in addition to and above mere "better than" judgments.**

Arguments Against all Standard Scales

- A standard utility representation requires transitivity of R.
- Hence, any argument against transitivity are arguments against standard scales.
- Semiorders / interval orders (Luce 1956; Tversky 1969; Fishburn 1991): Non-transitive equal goodness results from measurement errors and any indistinguishability for "equally good" comparisons. There is a utility representation but it's more complicated than the ones discussed so far.
- Non-transitive strict betterness (Temkin 1987, 2012; Rachels 1998, 2001): There are strong normative arguments against the transitivity of strict "better than" comparisons. Non-transitive value relations do not have any of the simple utility representations discussed so far.

▷ **There are both normative and epistemic arguments against the claim that all uses of "better than" have one of the standard scales discussed so far.**

My Suggestion

- Accept a ratio scale for aspects of uses of “better than” based on extensively measurable dimensions.
  Examples: costs, time, weight, ...
- By default assume ordinal scales or canonical, rank-based utilities for aspects of uses of “better than” that are not based on extensively measurable dimensions.
  Examples: value of personal relations (e.g. friendship), value of social systems (e.g. democracy), aesthetic value (e.g. of a painting)

▷ **Problem: The two cannot be combined! Should we bite the bullet and accept (fallacious) practical arguments?**

References

- Krantz, D. H., Luce, R.D, Suppes, P, Tversky, A. (1971, 1989, 1990): *Foundations of Measurement*, Vol. I-III. Academic Press.
- Lassiter, D. (2017): *Graded Modality: Qualitative and Quantitative Perspectives.* Oxford University Press.
- Rast, E. (2020): The multidimensional structure of 'better than'. *Axiomathes*, forthcoming, publ. online 2020. DOI 10.1007/s10516-020-09525-4.
- Rast, E. (2022): *Theory of Value Structure: From Values to Decisions.* Lexington.
- Roberts, F. (1979): *Measurement Theory.* Addison Wesley. Cit. in Roberts (1985).
- Soria-Ruiz, A. (2021): Value and Scale: Some Observations and a Proposal. *Organon F*, Vol. 28, No. 3, pp. 596-625.