Intersubjectivity of measurement across the sciences: Unit identification and dissemination

Andrew Maul
University of California, Santa Barbara

Luca Mari
Università Cattaneo

Mark Wilson
University of California, Berkeley

UCSB

Talk outline

- Some context: objectivity and intersubjectivity as features of measurement quality
- Intersubjectivity in the physical sciences
- Intersubjectivity in the human sciences
- Comparing the contexts

UCSB

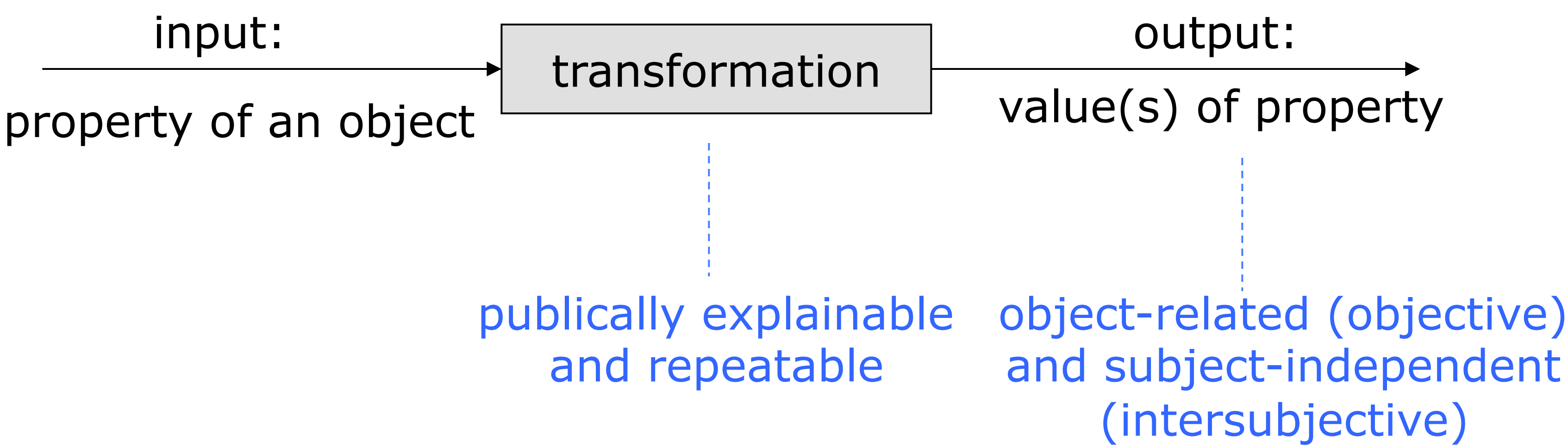
Context

- Luca Mari, Mark Wilson, and I are writing a book titled *Measurement Across the Sciences.*
- We hope to help build a shared concept system and vocabulary for measurement across the physical and human sciences.
- We propose that measurement is a process *of property evaluation* whose results are *credibly documented*.
- Two major conceptual dimensions of measurement quality in need of credible documentation:
  - *Object-relatedness* (“objectivity”)
  - *Subject-independence* (“intersubjectivity”)

Opening the black box of measurement

- A ”black box” characterization is insufficient to distinguish measurement from, say, guessing or opinion.
- Justifying the epistemic authority of measurement requires “opening the box” to identify what features of the measurement process guarantee the quality of the results.

Opening the black box of measurement



UCSB

Opening the black box of measurement

- *Objective*: specific to a given property of the object under measurement, thus independent of any other property of the object or surrounding environment
- *Intersubjective*: interpretable in the same way by different subjects in different places and times, because it is expressed in a form independent of the specific context and only refers to entities which are universally accessible

UCSB

Structure of the inference: ordinal example

Premise 0: hardness can be measured on an ordinal scale

Premise 1: $H[a]$ has been measured to be 7 on the scale s_1

Premise 2: $H[b]$ has been measured to be 8 on the scale s_2

Premise 3: $s_1 = s_2$

Conclusion: $H[b] > H[a]$

UCSB

Structure of the inference: ratio example

Premise 0: length can be measured on a ratio scale

Premise 1: $L[a]$ has been measured to be 1.23 times a given length unit u_1

Premise 2: $L[b]$ has been measured to be 2.34 times a given length unit u_2

Premise 3: $u_1 = u_2$

Conclusion: $L[b] = 2.34/1.23\ L[a]$

UCSB

Structure of the inference: general form

Premise 0: a given *property* can be measured on a given *scale type*

Premise 1: the *property* of object *a* has been measured to be v_1 with respect to the scale s_1

Premise 2: the *property* of object *b* has been measured to be v_2 with respect to the scale s_2

Premise 3: $s_1 = s_2$

Conclusion: the formal relation between v_1 and v_2 corresponds to the empirical relation between the properties of *a* and *b*

UCSB

Intersubjectivity across the sciences

- Premise 3 (i.e., $s_1 = s_2$) is an empirical hypothesis.
- Its justification establishes the metrological traceability of measurement results to a reference scale: the same result should refer to the same empirical situation independently of where and when and by whom it was obtained.
- This makes measurement results context-independent, and identically interpretable by different measurers.
- The claim that measurement is also possible in the human sciences is premised on the solution of this problem.

The traditional solution in physical measurement, part I

- Traditionally, a reference property (or properties) for a scale are identified as those realized by a given (set of) object(s), which we'll call s_0
  - For (general) quantities, the (individual) quantity realized by a given object defines the *unit*

- In the case of length:

$$\mathrm{m} := L[s_0],$$

where m is the unit (e.g., the metre) and s_0 is the primary standard (e.g., a given rod)

The traditional solution in physical measurement, part II

- If *a* and *b* can actually be compared to s_0, Premise 3 (i.e., $s_1 = s_2$) requires:
  - That s_0 was used in both comparisons, and
  - That s_0 did not change between the two comparisons

- More generally, we identify a sequence of measurement standards, $\langle s_1, s_2, \ldots \rangle$, such that s_{i+1} is accessible by s_i and $L[s_{i+1}]$ is guaranteed to be empirically indistinguishable from $L[s_i]$ (i.e., $L[s_{i+1}] \approx L[s_i]$)

UCSB

The newer solution in physical measurement

- The traditional solution is rooted on one (set of) object(s), which has drawbacks:
  - ownership confers control over the whole system
  - the stability of the whole system depends on the stability of the primary set/standard

- The newer solution: identify the reference properties as those realized by a phenomenon (or class of objects), thought to be stable according to the best available theory

- Again in the case of length:

$$\text{m} := L[P_C]$$

where P_C is, e.g., a beam of light in vacuum for a given duration

UCSB

Underlying principles

i. identifying a reference scale and using it are distinct processes;

ii. measuring instruments are designed and operated under the assumption that the reference scale for the property they measure is already and independently defined;

iii. measuring instruments are designed and operated so as to compare the property to be measured and the reference scale: the property to be measured is a property of the object under measurement; the reference scale is either conveyed by some measurement standards (in direct synchronous methods of measurement) or are stored in the instrument via its calibration (in direct asynchronous methods of measurement).

Structure of the inference: chess-playing ability

Premise 0: chess-playing ability can be measured on an ordinal scale

Premise 1: $C[a]$ has been measured to be 7 on the scale s_1

Premise 2: $C[b]$ has been measured to be 8 on the scale s_2

Premise 3: $s_1 = s_2$

Conclusion: then $C[b] > C[a]$

Analogue of the traditional solution in psychosocial measurement

- We may identify a set of progressively more skilled individuals as references (s_0)

- A new individual may enter into competition with these individuals to determine where in the sequence (s)he should be located

- As before, this depends on
  - the empirical accessibility of s_0
  - the stability of the property in s_0

Analogue of the newer solution in psychosocial measurement, part I

- Just as in physical measurement, we could identify the reference properties as those realized by a class of persons, in given conditions, thought to be stable according to the best available theory

- Example: the behavior of the reference set could be encoded and stored as a set of “prototypical virtual chess players”, thus avoiding the problems of empirical accessibility and stability:
  - these entities are now informational rather than physical, and therefore could be easily disseminated
  - algorithms are by definition stable

UCSB

Analogue of the newer solution in psychosocial measurement, part II

- In psychosocial measurement, an individual is often asked to react to a set of specific, constrained challenges or prompts ("items")

- Ideally, variation in responses to items is caused by variation in the property intended to be measured (*objectivity*)
  - This relationship is usually assumed to be probabilistic, and can be formalized e.g. via the Rasch model:

$$\text{log-odds}[X_{ni} = 1] = \theta_n - \delta_i$$

where θ represents the property of the person and δ represents the severity of the item

Structure of the inference: reading comprehension ability

Premise 0: reading comprehension ability can be measured on an interval scale

Premise 1: $R[a]$ has been measured to be 1.23 times a given RC unit u_1

Premise 2: $R[b]$ has been measured to be 2.34 times a given RC unit u_2

Premise 3: $\mathrm{u}_1 = \mathrm{u}_2$

Conclusion: then $R[b] = R[a] + (2.34 - 1.23)\mathrm{u}$

Analogue of the newer solution in psychosocial measurement, part III

- If we supposed that:
  1) reading comprehension ability can be measured on an interval scale, and
  2) is assessed via responses to a set of items, and
  3) item response data from the administration of these items to individuals in the relevant population is dependably found to fit a Rasch model, then:

- Each value of θ could be considered an instantiation of a “virtual person”, with a given profile of probabilities of success on all items, and

- A unit can then be defined as the distance between any two arbitrarily-chosen values of θ (or δ)

UCSB

Analogue of the newer solution in psychosocial measurement, part IV

1. The solution just described yields a unit specific to a given instrument (set of items)
   - Premise 3 is satisfied only for comparisons made using that instrument

2. If a large number of items are found to dependably fit a Rasch model, new instruments can be created using any sub-sample of these items
   - Premise 3 is now satisfied for comparisons made using different instruments derived from the item bank

3. Other instruments could be made traceable to the same reference objects, following the same procedure
   - Premise 3 is now satisfied for comparisons made using different instruments calibrated to the same reference scale

Additional issues in the human sciences

- While Premise 0 seemed unproblematic in the physical cases, it is not so in the psychosocial cases
  - what does it mean to assert that reading comprehension ability can be measured on an interval scale?

- In addition to stability of the primary reference set/ standard, there may be concern about stability of the general property definition itself
  - has the definition of reading comprehension ability changed in the last 20 years?

Conclusions

- Traceability to a reference set/standard can, at least in principle, be structurally guaranteed in both physical and non-physical measurement.

- The human sciences seem to be behind the physical sciences in terms of recognizing the importance of this task.
  - Is solution will depend on clarification of property definitions and theories of measurement as well as traceability chains.

- We hope that presenting the requirements for measurement quality in a single and consistent framework will help facilitate progress in our field.

Thank you for your time!

amaul@ucsb.edu