Measurement xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Intersubjectivity of measurement across the sciences

Andrew Maul [a,*], Luca Mari [b], Mark Wilson [c]

[a] Gevirtz Graduate School of Education, University of California, Santa Barbara, CA USA
[b] School of Industrial Engineering, Università Cattaneo – LIUC, Castellanza, VA, Italy
[c] Graduate School of Education, University of California, Berkeley, CA, USA

ARTICLE INFO

Article history:
Received 16 August 2018
Received in revised form 29 August 2018
Accepted 30 August 2018
Available online xxxx

Keywords:
Measurement
Philosophy of measurement
Intersubjectivity
Units
Measurement standards
Reference properties

ABSTRACT

A critical condition for the quality of measurement results is that they be interpretable in the same way by everyone, even though they may have been obtained in different contexts by different individuals using different instruments: in other words, they should be subject-independent, or intersubjective. For both physical properties and psychosocial properties, intersubjectivity can be secured by establishing the metrological traceability of the measurement results to a measurement unit, and more generally to a set of reference properties, though at present such solutions are less commonly found in psychosocial applications. In this paper we describe traditional and newer solutions to the problem of intersubjectivity in the physical sciences, and then explore how these and other solutions can apply to non-physical measurement as well. The fact that, despite their differences, the metrological traceability to references can be structurally guaranteed in both physical and non-physical measurement and can be presented in a single and consistent framework is a significant step towards the development of a conception of measurement across the sciences.

1. Introduction

We begin by asking the reader to imagine the following scenario. Suppose that two individuals, in two different places and times, using two different instruments, each measure the hardness (considered as an ordinal property) of two different objects. That the first individual finds the hardness of her object to be "8", and the second individual discovers the hardness of his object to be "7"; both individuals then report their results, referring to the same scale. Under what conditions would we be able to confidently conclude that the hardness of the first object is actually greater than the hardness of the second object?

We might also imagine a second scenario, identical to the first in structure, except that instead of measuring the hardness of two physical objects, we were instead concerned with evaluating the chess playing ability of two individuals. Again we could ask the question: under what conditions would we be able to confidently conclude that the chess playing ability of the first person is actually greater than the chess playing ability of the second person?

Further, does anything about the answers to these questions change if we instead consider properties commonly modeled as quantitative rather than ordinal, such as length, temperature, or reading comprehension ability [18]?

The scenarios described here highlight a critical issue regarding the quality of measurement results, and both pose a special case of a more general question: under what conditions are we willing to accept that measurement results are of high enough quality to be depended upon? As we have argued elsewhere [17], the societal role of measurement is underwritten by trust in the quality of the information obtained. In particular, we proposed that there are two major dimensions of measurement quality in need of credible documentation: object-relatedness, or *objectivity*, and subject-independence, or *intersubjectivity*. From our perspective, then, the primary task facing anyone wishing to justify the dependability of measurement results—regardless of whether the measurement is of a physical property such as hardness or a non-physical property such as chess playing ability—is documentation of the structural features of the measurement process that serve to secure both objectivity and intersubjectivity.

To perhaps state the obvious, this is no trivial task. The framing of the two scenarios just presented was intended to call attention in particular to one aspect of the overall task, related to the credible documentation of the features of the measurement system that allow measurement results obtained by different individuals in different times and locations and using different instruments to be meaningfully compared—in other words, to be subject-

* Corresponding author.
E-mail address: amaul@ucsb.edu (A. Maul).

https://doi.org/10.1016/j.measurement.2018.08.068
0263-2241/

independent, or intersubjective. In order to focus on the requirement of intersubjectivity and the ways in which it may be secured, we assume here that objectivity has already been secured—in other words, we assume that the information obtained on hardness and chess playing ability (and other examples used in this paper) truly does relate to that property of the object under measurement, which in turn requires the assumptions that (a) the properties in question have been sufficiently well-defined, making definitional uncertainty (JCGM, [11]: 2.27) negligible, and (b) that the measurement instruments were not sensitive to other properties (also termed "influence properties"), making instrumental uncertainty negligible. These are clearly not trivial assumptions. For now, though, committing to them allows us to focus on the following two primary questions:

(1) How can intersubjectivity be structurally guaranteed?, and
(2) How does the answer to the first question change, depending on the area of application (e.g., physical properties, psychosocial properties, etc.)?

This paper is structured as follows. In Section 2 we first say more about the problem itself, to attempt to clarify what is at stake. We then briefly review in Section 3 solutions found in physical metrology to guarantee intersubjectivity, followed in Section 4 by a similar review of solutions that have been or could be proposed in the human sciences. We conclude with a discussion of the similarities and differences of these solutions. We note some issues that condition the possibility of guaranteeing intersubjectivity in the human sciences, but argue that, despite their differences, traceability can in principle be structurally guaranteed in both physical and non-physical measurement settings and can be presented in a single and consistent framework, which we propose is a significant step towards the development of a single, consistent conception of measurement across the sciences.[1]

2. The problem to be solved, the conditions to be guaranteed

Measurement results convey information on the relation between the measurand (JCGM, [11]: 2.3) and one or more reference properties. In the canonical case of additive quantities, this relation is between the measurand and another quantity of the same kind, chosen as the quantity unit. By stating that, for example, the length $L[a]$ of rod a is 1.23 m, possibly together with some measurement uncertainty, we claim that the two lengths identified as $L[a]$ and the metre have been compared and the former is 1.23 times greater than the latter, i.e., that $L[a]/\mathrm{m} = 1.23$. Property evaluations reported in reference to interval, ordinal, and even nominal scales[2] are analogous: their results are relational, again involving a measurand and other reference properties; only the structure of the involved relation changes. For example, the result that the temperature $T[a]$ of rod a is 23.4 °C means that both a unit and a zero temperature have been conventionally chosen – let us designate them as u_C and z_C respectively – and $(T[a] - z_C)/(u_C - z_C) = 23.4$. Likewise, the result that the hardness $H[a]$ of a is 7 in the Mohs scale conveys the information that the hardness of a is equivalent to that of the seventh element of the scale, i.e., quartz.

In all of these cases, before the measurement a set of reference properties or quantities is established (and the scale type determines how this can be done), and the measurement then consists of the (direct or indirect, explicit or implicit) comparison of the measurand with the elements of the set. The fact that measurement results are in this sense relational has the consequence that properties of objects become comparable not only by empirically comparing objects by their properties but also mathematically via the values of these properties. For example, if $L[a] = 1.23$ m and $L[b] = 2.34$ m, then the ratio $L[b]/L[a]$ is scale invariant, and therefore we can infer that $L[b] = 2.34/1.23\ L[a]$, even without directly comparing a and b by their lengths. Analogous conclusions can be drawn in the interval and ordinal cases mentioned above: if $T[a] = 23.4$ °C and $T[b] = 34.5$ °C, then the ratio $(T[b]-z_C)/(T[a]-z_C)$ is scale invariant, and therefore we can infer that $T[b] = 34.5/23.4\ (T[a]-z_C) + z_C$, even without directly comparing a and b by their temperatures; and if $H[a] = 7$ Mohs and $H[b] = 8$ Mohs, then the relation $H[b] > H[a]$ is scale invariant, and therefore we can infer that $H[b] > H[a]$, even without directly comparing a and b by their hardnesses.

The validity of these inferences depends on a number of premises. First, as stated previously, we assume that the measurement results are objective, in the sense of referring to the property in question (hardness, temperature, length) and not to anything else. This also involves the condition that it is in fact possible to consistently measure the considered properties on the assumed scale types (and therefore that lengths can be measured on a ratio scale, and so on), a nontrivial premise that we will not comment further on here (though see [14]). A second premise, more hidden but not less critical (and in fact the one we mainly explore here), is that the scale against which the two measurands were compared (i.e., the scale generated by the metre in the ratio example) is the same. It should be noted that this is unrelated to the linguistic choice of using the same term ("metre" or whatever else) for referring to the scale in the two measurement results: since quantity units are quantities in turn, and more generally reference properties are properties in turn, this is an empirical hypothesis and as such it must be justified. The actual inference therefore has this structure[3]:

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[4] between v_1 and v_2 corresponds to the empirical relation between the properties of a and b

[1] The more general issue of the comparability of measurement concepts and practices across different areas of application has been the subject of a significant amount of scholarship over the past century. This is a complex subject on which one of the authors co-organized the 2016 Joint IMEKO TC1-TC7-TC13 Symposium, "Metrology Across the Sciences: Wishful Thinking?", 3–5 August 2016, Berkeley, USA, whose proceedings have been published in the IOP Journal of Physics: Conference Series, 772, 2016. Some other volumes that could be usefully considered on this matter are, e.g., Berglund et al. [1], Boumans [2], Schlaudt and Huber [19].

[2] By "scale" (or "reference scale"), we mean an appropriate set of individual properties that are identified as properties of objects, not a set of values of a property. (Values of a property result from the scale, not vice versa: trivially, we first define the metre then we assign the value 1 m to it). Also, following the tradition of Stevens [20], we use the terms "scale" and "scale type" even for nominal properties, for which the concept 'scale', which recalls ordering, would not be applicable. We accept here that measurability is not a priori constrained by scale type. A justification of this position is given in Mari, Maul, Torres Irribarra & Wilson [14].

[3] Interestingly, the structure of this inference is the inverse of the structure of measurement according to the representational theory of measurement [12]: here the empirical relation of properties is the result of sufficiently intersubjective measurements; in RTM it is a condition to construct measurements.

[4] It could be noted that there is another trivial premise at work in these examples, which is that the numbers appearing in the measurement results have formal properties of their own: e.g., in the ordinal cases, that 8 > 7; in the ratio and interval cases, that two numbers have a given ratio, etc. In general such a premise need not be stated, but making it explicit helps make clear the idea that the conclusion is the (non-trivial) empirical analogue of the (trivial) formal relation between the numbers involved.

Please cite this article in press as: A. Maul et al., Intersubjectivity of measurement across the sciences, Measurement (2018), https://doi.org/10.1016/j.measurement.2018.08.068