Measurement xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Measurement

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

# The quality of measurement results in terms of the structural features of the measurement process

Andrew Maul[a,*], Luca Mari[b,1], David Torres Irribarra[c], Mark Wilson[d]

[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] Escuela de Psicología, Pontificia Universidad Católica de Chile, Chile
[d] Graduate School of Education, University of California, Berkeley, CA, USA

ARTICLE INFO

*Keywords:*
Measurement
Philosophy of measurement
Measure
Quality of measurement

ABSTRACT

In both scientific and lay settings, measurement is considered a privileged source of high-quality information, and is commonly associated with precision, accuracy, and dependability. However, it is not always clear what features of the measurement process justify this public trust, and how the quality of measurement results in different domains of inquiry can be compared. In this paper, we first argue that the quality of measurement results depends on their object-relatedness (“objectivity”) and subject-independence (“intersubjectivity”) and is justified on the basis of the structural features of the measurement process, as well as features of the inputs or the outputs of the process. Given this perspective, we analyze three general measurement methods, according to which a measurement process can be structured and performed, which may be called (a) direct synchronous, (b) direct asynchronous, and (c) indirect. In addition to the value of these distinctions for the process of designing measuring instruments, they allow us to highlight the different roles of models, theories, and computations in measurement. We then attempt to apply this classification strategy in the context of the social sciences by discussing the role of (1) the definition of the measurand and (2) the theory connecting the measurand to the measurement results in each of these measurement methods, and how they can or cannot be conceptualized from the perspective of measurement theories in the social sciences. This leads us to the conclusion that the differences between physical and non-physical measurement are historical and contextual rather than essential; that is, in both cases, the quality of measurement results can be effectively evaluated from a structural perspective.

## 1. Introduction

It would be difficult to overstate the value and importance of measurement in nearly every aspect of society. Every time we eat food, take prescribed medicine, fly in an airplane, use a cell phone, or step inside a building we place our trust in the results of measurements – and, for the most part, that trust seems well-earned, and as such measurement is commonly associated with precision, accuracy, and objectivity [33]. Against this backdrop, it seems little wonder that the social sciences (including psychology, sociology, economics, and field-specific areas of research, such as education) have, since their inception, attempted to incorporate measurement into their activities as well. However, despite – or perhaps, to at least some extent, because of – the ubiquity of measurement-related concepts and discourse, there remains a remarkable lack of shared understanding of these concepts across (and often within) different fields, perhaps most visibly reflected in the vast array of proposed definitions of measurement itself (see the review and related discussion in [20]). In addition to obviously hampering communication across different disciplinary fields regarding shared methodological principles, such a lack of common understanding hints at the possibility that the same terms – “measurement”, “measurement result”, “measurement model”, etc. – are used with very different and possibly even incompatible meanings, with potentially disastrous results. [2]

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

[1] One of the authors is a member of the Joint Committee on Guides in Metrology (JCGM) Working Group 2 (VIM). The opinion expressed in this paper does not necessarily represent the view of this Working Group.

[2] The comparability of measurement concepts and practices across different scientific disciplines has been the subject of a significant amount of scholarship over the past century. The present paper is aimed at contributing to the general endeavor of improving the understanding of measurement across the sciences, 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., [2], [5], [6], [40]

http://dx.doi.org/10.1016/j.measurement.2017.08.046
Received 6 May 2017; Received in revised form 12 July 2017; Accepted 28 August 2017

It would seem, then, that the clarification of foundational measurement concepts should (continue to) to be a high priority: in terms not only of the definition of measurement itself, but also of the identification of those features of measurement that justify its commonly-afforded degree of public trust and social prestige. Justification of the dependability of measurement results, in turn, depends on identifying those features of the measurement process that ensure (or, at least, confer high likelihood upon) the quality of the results.

There are at least two (categories of) reasons why measurement-related concepts have become so difficult to define in a consistent way across different fields. First, as the scope of measurement has broadened, it is not always obvious what – if indeed anything – is common among all the processes claimed to be measurements, but surely a shared body of knowledge cannot be found in the technical details on which measurement science advances in each specific field. Second, the scholarly treatment of the concept of measurement has focused since the second half of the 20th century on purely formal criteria, thus abstracting from the concrete realization of the process, up to the point that one of the reference books on representational theories of measurement is titled "Abstract measurement theory" [28], and other researchers in the field have made claims such as that "we are not interested in a measuring apparatus and in the interaction between the apparatus and the objects being measured. Rather, we attempt to describe how to put measurement on a firm, well-defined foundation" [36] and "The theory of measurement is difficult enough without bringing in the theory of making measurements" [18]. This emphasis on a formal characterization of measurement is consistent with the expansion of measurement into many new domains of application, abandoning definitions that could be tied to requirements of specific areas; abandoning for instance elements tied to the traditional realization of measuring systems operating on the basis of transductions implemented by physical sensors possibly due to the fact that the evaluation of non-physical properties[3] cannot conform to it. As a consequence, theoretical interpretations of measurement have become so abstract that they may be unable to provide a convincing and useful demarcation of measurement from formally similar processes that are generally thought to lack epistemic authority, such as most instances of the expression of subjective judgments and opinions (as already acknowledged, e.g., by [37]: "In the social sciences, in particular, most evaluations are not measure[ment]s, but rather mixtures of opinion and estimation.")

One may question whether working on the definition of 'measurement' is a worthwhile endeavor. Here our position on this matter is also practical: there is a social interest in sharing scientific and technical vocabulary across disciplines[4], particularly in the case of an infrastructural activity like measurement [15], and there is a social acknowledgment of the epistemic authority of measurement, which has critical consequences in particular in terms of public trust attributed to the outcomes of putative measurement processes and the resources devoted to such processes. If the idea of "measurement" can be invoked at will, without understanding or concern for what has historically made it a valued practice, it becomes simply a rhetorical device, risking to discredit its practice in general.

We propose here is that measurement is a process characterized by its *structure*, not only by the specification of the functional relationship connecting its inputs to its outputs: what is required is an explanation of *how* the process does what it does, not only of *what* it does. While, for example, measurements based on thermal expansion thermometers and on electrical resistance thermometers could be treated as interchangeable in functional terms, they clearly have different structures: even if the function is the same, its implementation/realization is distinct. Whereas a functional relationship relies solely on a black box model, a structural model involves identification of the invariant aspects that are implemented in the experimental process – and this, in turn, as we will argue, is what provides justification of the claim that measurement results are publicly trustworthy. As a corollary, any purely black-box (meta-)model cannot adequately account for relevant features of measurement, and thus is not sufficient for the purpose of understanding the quality of measurement results.

In the metrological tradition the general description of the structure of a measurement is provided by a so-called "measurement method", the "generic description of a logical organization of operations used in a measurement" according to the VIM [[15], def. 2.5]. This paper proposes some preliminary considerations and examples to show that different measurement methods, each with their own specific structures, share the same invariant meta-structure (on the concept of measurement meta-structure see also [22], that the present paper expands). With some provisos – including the availability of a sufficiently well-detailed definition of the general property of which the measurand is an instance – this invariance is independent of the nature of the measurand and therefore spans the measurement of both physical and non-physical properties.

The next section is devoted to introducing this meta-structural understanding of measurement in reference to three basic measurement methods, as developed in metrology, and to discussing the conditions for the quality of measurement for each method. On this basis, Section 3 explores how these structures apply in the case of non-physical properties, and argues that the most critical barrier to understanding the operative structure of non-physical measurement processes – and, therefore, to understanding how the dependability of such measurement results is justified – relates not to any fundamental distinction between the two areas, but to the often imprecise way in which general non-physical properties are defined.

## 2. A meta-structural understanding of measurement

### *2.1. Black-box characterizations of measurement*

Under the general hypothesis that measurement is a process that operates on inputs (at least the measurand, in the case of direct measurement methods[5]) to produce outputs (at least the measurement result), measurement could be characterized as an instance of the black box meta-model that describes processes as entities that transform inputs to outputs (Pane [a] in Fig. 1).

Conventionally, measurement is a process aimed at producing information in the form of values (e.g., 0.1234 m) attributed to properties (usually, quantities) of objects (e.g., the length of a given object). However, such a characterization is not specific to measurement: other processes, such as, say, quantitative guessing, take as input the property of an object and produce in output one or more values that are attributed to the property. Let us call "property evaluation", or simply "evaluation" for short, *any* process with this black box characterization

[3] For the sake of generality, the term "property" is used rather than "quantity" throughout this paper. The *International Vocabulary of Metrology* (VIM) defines quantities as specific kinds of properties [[16], def. 1.1].

[4] A basic reason for the complexity of this endeavor is the (usually unavoidable and in fact appropriate) specialization of the scientific and technical disciplines, which triggers the construction of specific terminologies. An interesting example of an attempt to overcome lexical hyper-specialization while maintaining scientific and technical correctness is Electropedia, "the world's most comprehensive online electrical and electronic terminology database containing more than 20,000 terms and definitions", that makes the series of standards IEC 60050 freely accessible online at www.electropedia.org.

[5] The possibility of 'direct' measurement (i.e., more correctly, direct *methods* of measurement), is sometimes dismissed as naive, via the argument that "all measurements are indirect in one sense or another" because "not even simple physical measurements are direct" given that, e.g., "the physical weight of an object is customarily determined by watching a pointer on a scale. No one could truthfully say that he 'saw' the weight" [Guilford 1936]. Of course, this is not the meaning assumed here and, e.g., by the VIM [[15], def. 2.5 Note], which notes that measurement methods are either direct or indirect. In this view, direct methods are simply those in which the measuring instrument directly interacts with the object under measurement.