

PAPER • OPEN ACCESS

Can there be one meaning of “measurement” across the sciences?

To cite this article: Luca Mari *et al* 2019 *J. Phys.: Conf. Ser.* **1379** 012022

View the [article online](#) for updates and enhancements.

You may also like

- [The Fundamental Relation between Halo Mass and Galaxy Group Properties](#)
Zhong-Yi Man, Ying-Jie Peng, Jing-Jing Shi et al.
- [First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole](#)
The Event Horizon Telescope Collaboration, Kazunori Akiyama, Antxon Alberdi et al.
- [An Electrochemical Impedance Study on the Influence of Flow on the Corrosion Inhibition of Carbon Steel by Fatty Amines in Association with Phosphonocarboxylic Acid Salts](#)
Nadine Pebere, Nathalie Ochoa and Bernard Tribollet



DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Can there be one meaning of “measurement” across the sciences?

Luca Mari¹, Andrew Maul², Mark Wilson³

¹ School of Industrial Engineering, Università Cattaneo, Castellanza, VA, Italy

² Gevirtz Graduate School of Education, Univ. of California, Santa Barbara, CA, USA

³ Graduate School of Education, University of California, Berkeley, CA, USA

E-mail: lmari@liuc.it

Abstract. The *International Vocabulary of Metrology* (VIM) gives the following definition of ‘measurement’: “process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity”. In pursuit of a common understanding of measurement, we explore this condition of *reasonableness* across the sciences, asking in particular whether there are common features of the structure of measurements in different fields of study that could meet this criterion. We propose that a general sufficient condition for measurement is that it is an experimental process of evaluation of empirical properties that produces *explicitly justifiable information*, which, we argue, is consistent with characterizations of science and epistemology more generally.

1. Introduction

The concept of measurement, along with associated concepts and terminology, is frequently invoked in many different scientific fields as well as throughout society at large, and the range of activities referred to as measurements or as being related to measurements continues to expand and diversify. In many ways this trend is simply a corollary to the continued expansion of the scope of science itself, given how integral measurement is, and is perceived to be, to many aspects of scientific inquiry [1]. Throughout the twentieth century and up to the present moment, a great deal of time and attention has been given by philosophers of science and the general public to the question of *what science is*, as distinct from other types of human activities (see, e.g., [2]); as could be expected, the similar and more local question of *what measurement is* has also received significant attention, particularly in light of claims made by psychologists and other human scientists starting around the turn of the twentieth century regarding the measurability of psychological and social properties (see, e.g., [3]).

Although an observer might note that identifying such distinct endeavors with the same terms (“measurement”, “measurability”, etc.) could be merely a lexical peculiarity based on some homonymy, we contend that significantly more is at stake: just as debates regarding the demarcation between science and pseudo-science were given urgency insofar as they bore on the degree of public trust that should be afforded to fields such as psychoanalysis and Marxist economic theory, debates regarding the legitimacy of claims about measurement similarly bear on the status of the human sciences, with some going so far as to assert that such claims are not only illegitimate but “pathological” [4] insofar as they misunderstand the necessary and sufficient conditions for an activity to properly be characterized as a measurement. More positively, we hypothesize that an exploration of the interdisciplinary similarities and differences in approaches to measurement (and discourse about measurement) may shed new light on the question



of what measurement is, and what features of the measurement process serve to warrant its commonly acknowledged trustworthiness.

2. Can there be one meaning of “measurement” across the sciences?

The *International Vocabulary of Metrology* (VIM) [5], possibly the most authoritative existing source regarding measurement terms and concepts, gives the following definition of ‘measurement’: “process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity”. This is, in essence, a black-box characterization of measurement: measurement is defined as an experimental process whose input is a quantity, and whose output is one or more quantity values. (Two notes. First, the term “experimental” plausibly should be understood in a generic sense here, in contrast to “theoretical” (e.g., as based on a thought experiment); in other words, measurements are processes that interact with empirical features of the world. Second, we do not consider here the possible generalization to non-quantitative properties of objects, i.e., whether non-quantitative properties are also measurable – though see [6].)

Read at face value, this definition would equally apply to processes such as guessing or statement of personal opinion, if these were phrased in quantitative terms (e.g., “I think it will be three degrees warmer tomorrow than it is today”; “I’m ten percent happier today than I was yesterday”), with only the condition that some experimental activity was performed for obtaining the outcome. Given that such processes are not generally considered measurements, the validity of the VIM’s definition critically relies on the adverb “reasonably”.

Clearly reasonableness is not a formal condition; rather, it can be considered as a means to provide an encompassing definition, acceptable by the broadest set of researchers, independently of their field of activity (indeed, who would refuse to accept that measurement should produce “reasonable” results?). But this condition is not sufficiently specific: it could be that, for example, some guesses about empirical quantities could be considered reasonable, but nevertheless we might refuse to accept them as measurements. Hence the problem we explore here is: *can there be one criterion of “reasonableness” that is sufficient to characterize all processes that across the sciences are commonly considered measurements?*

This key problem should be interpreted as related to a descriptive, not normative, characterization of measurement: “measurement” is not a trademarked term, and in principle everyone is free to call “measurement” whatever (s)he likes. But this does not take into account the special epistemic role commonly acknowledged to measurement: were it the case that measurement and, e.g., opinion-making were equally valuable, why should money be spent for buying, maintaining, and operating expensive measuring instruments?

It is a critical condition for the criterion of “reasonableness” we are seeking that it be one that applies across the sciences. A *specific* sufficient condition could be, in particular, that the process is realized by means of a properly operated calibrated physical transducer that is sensitive to the property intended to be measured – which is indeed the sufficient condition usually applied to characterize physical measurements: it works for physical properties, and may also provide some guidelines for designing measurement systems for psycho-social properties. But treating it as a general sufficient condition would correspond to an old-fashioned physicalism, which assumes that only physical properties are measurable, thus considering non-physical properties as unmeasurable simply *by fiat*. Hence any such candidate criterion should be independent of domain-specific strategies of design of measuring instruments and implementation of measurement processes.

3. Different subject matters, different processes...

In the physical sciences, the properties intended to be measured are often already embedded in a well-established body of knowledge, and the work is usually aimed at the refinement of existing instruments or mathematical techniques. The term “instrumentation and measurement” (as found, for example, in the title of the IEEE journal *Transactions on Instrumentation and Measurement*) clearly conveys this message, and the books on measurement systems in these fields are largely devoted to presentation of

the features and modes of operation of measuring instruments (e.g., [7, 8]). In the human sciences, such quantitative knowledge is less developed, and thus activities aimed at measurement necessitate first of all the task of formally characterizing the properties of interest, in such a way that quantitative information on them can be acquired and then mathematically processed. The key issues in these cases are the meaningfulness (e.g., [9]) and more generally the validity (e.g., [10]) of the hypotheses upon which the formalization is based, and such issues are the focus of measurement sub-disciplines within specific human sciences such as psychometrics and econometrics, as components of psychology and economics respectively.

This multiplicity of perspectives reflects the fact that researchers operating in different fields face quite different problems and challenges, and will thus have different associations with the concept of measurement. However, this does not preclude the possibility that there are common elements of the structures of the processes referred to as measurements in different fields, and that identifying these common elements may suggest paths towards the formulation of conception of measurement applicable across the sciences.

4. ... with some structural commonalities...

In this analysis we draw from previous works, in which we have discussed the very definition of ‘measurement’ (e.g., [11]) and its epistemology (e.g., [12]), the stereotypes that surround measurement (e.g., [13]), and in particular the mistaken assumption that measurement is identical to quantification [6]. A background commonality that emerges from these studies is that measurement (1) is an experimental process that (2) produces information on empirical properties of objects (where objects can be bodies, systems, phenomena, events, processes, individuals, organizations, etc.), such information being (3) in the form of values of property. This is consistent with the VIM’s definition quoted previously and, as with the VIM’s definition, it is an abstract, black-box (i.e., input-output) characterization of measurement, that nevertheless serves to delimit the scope, in particular by distinguishing measurement and computation (where the latter produces information from other information).

Even without “opening the box”, such a characterization already includes several key conditions, related to

(i) *the procedural nature of the process* (highlighting that measurement is designed on purpose and therefore justifying the focus on the experimental and mathematical instrumentation as a tool to implement the procedure, as mentioned in the first category of cases above, more commonly emphasized in physical measurement), and to

(ii) *the definition of the property* intended to be measured and to which the produced information is attributed (highlighting that measurement requires modeling stages that are preliminary to data acquisition, as mentioned in the second category of cases above, more commonly emphasized in psycho-social measurement).

These can be interpreted as *necessary* conditions for a process of evaluation of an empirical property to be considered a measurement (an additional condition sometimes considered necessary is that the measured properties are quantities – this is in particular the VIM’s position, reflecting a long tradition in physical sciences; as previously mentioned, we do not develop this subject here). But of course the critical issue is about the *sufficient* conditions.

5. ... and a common emphasis on trustworthiness...

As was previously observed, measurement is usually associated with trustworthiness. Elsewhere (e.g., [14]) we have argued that the trustworthiness of measurement results (more or less equivalently for this concept we used also the term “dependability”) requires that they convey information that is *object-related* and *subject-independent*, or, more explicitly:

1) information that is specific to the measurand and, therefore, to a given property of the object under measurement. This means that the provided information should be independent of any other property of the object or the surrounding environment, which includes both the measuring system and the subject

who is measuring. This corresponds to guaranteeing that measurement results actually provide information about the measurand and not some other property. This condition is about the appropriate attribution of information to its claimed object: hence, it is a requirement of object-relatedness, or “objectivity” for short;

2) information interpretable in the same way by different users in different places and times. This corresponds to guaranteeing that measurement results are expressed in a form independent of the specific context and only referring to entities which are universally accessible, so that the meaning of a measurement result is unambiguous and can be easily reconstructed in principle by anyone, possibly on the basis of suitable conventions: hence, it is a requirement of subject-independence, or “intersubjectivity” for short.

According to this characterization, objectivity and intersubjectivity are independent features – something can be objective but not intersubjective, and vice versa –, that identify the two dimensions of measurement: the claim of the possibility of *obtaining information about empirical properties*, and the claim of the possibility of *socially reporting such information* (interestingly, these dimensions correspond to the two main stages of a measurement process, which Roman Morawski calls “conversion” and “reconstruction” [15] and Giovanni Battista Rossi and Francesco Crenna call “observation” and “restitution” [16]). It is thus through their objectivity and intersubjectivity that measurement results are considered trustworthy.

On the other hand, as intended here objectivity and intersubjectivity are not Boolean (i.e., yes-no) conditions: something can be more or less objective and more or less intersubjective. Hence for objectivity and intersubjectivity to be the sufficient conditions that we are looking for to characterize measurement as specific kind of evaluation, a threshold of minimum objectivity and intersubjectivity should be set. Thus the term “reasonable” in the VIM’s definition of ‘measurement’ once again does the heavy lifting: a given attribution of values would be considered reasonable (and therefore, would be considered a measurement) if both its objectivity and intersubjectivity were sufficient for the intended purpose for which the instrument was designed and performed. This highlights the pragmatic nature of measurement: the same measurement may be considered good for some purposes and bad for some others.

Emphasis on sufficient objectivity and intersubjectivity for a given purpose is then operatively useful, for the general guidelines it provides regarding the design and performance of measurements (e.g., in [17]) – though more work should be done to develop better guidelines – but it is still too specific at least in one respect: it would assume that measurement is always *good* measurement. While pragmatically this is sound – if we know that what we are doing is a bad measurement (e.g., with measurement uncertainty greater than target uncertainty, so that the produced information is useless) we avoid doing it – the concept ‘bad measurement’ as such is not contradictory, and bad measurements do not fulfill the condition of sufficient objectivity and intersubjectivity. In other words, in order to maintain the VIM’s characterization of “reasonableness”, objectivity and intersubjectivity are useful but still not sufficient: an even more general condition has to be identified.

6. ... and a focus on producing explicitly justifiable information

Science aims at the development of knowledge, where, following Plato, knowledge is commonly understood as *justified true belief*. But although science *aims* at truth, it cannot *guarantee* it, as plainly illustrated by its history, and thus the truth or falsity (and more generally the quality, were a criterion of quality given) of any given scientific theory cannot be a definitional component of what makes it scientific: a theory can be of low quality, and even eventually admitted to be false, and nevertheless can be scientific. Although discussions of the definition and essential features of science are still ongoing, it is generally agreed that *scientific theories must be explicitly justifiable*, in that their logical and evidentiary grounding must be clear and publicly evaluable [2]. Thus, *science is characterized by its structure rather than its outcomes*; it is no contradiction to say that a theory is both scientific and false, but it would be a contradiction to say that a theory is both scientific and untestable, even if it were true. Similarly, the trustworthiness of measurement results is not earned solely by their objectivity and

intersubjectivity, but by the *justification* that is provided for such, in the sense that an explanation can be given for how the structure of the measurement process serves to secure the trustworthiness of its results. The availability of such an explanation should, at least in principle, make the results reproducible, thus making possible their social critical analysis. Again, it is no contradiction to say that a given process is both a measurement and produces results that are of low quality – as in the case of a bad measurement – but it would be a contradiction to say that a process is both a measurement and cannot be justified, even if the results were (perhaps accidentally) accurate.

Of course, in many cases, particularly of non-scientific measurements, measuring systems remain black boxes and no justifications of their results are reported: what is important is that a justification *can* be reported, whenever required. Thus the idea of the general sufficient condition that we are seeking, applicable across the sciences, is that measurement is an experimental process of evaluation of empirical properties that produces *explicitly justifiable information*. In this way, the expectations of good measurement processes are continuous with the expectations of other sources of knowledge, in Plato's sense of knowledge as justified true belief. This is the case regardless of the subject matter and the details of the procedures involved in measurement.

The requirement that measurement results be explicitly justifiable helps explain why measurement cannot be adequately characterized solely using a black-box model: if a given attribution of value(s) to a property is claimed to be a measurement result (instead of, e.g., once again, a guess), it must be possible to explain how it was obtained, by “opening the box” and identifying the features of the process that secure the quality of the results. It is, of course, no simple matter to describe how such justification should take place in any given instance of measurement, and it is here in which we find significant differences across disciplines. As previously noted, the key activities associated with designing and using measuring instruments in the physical and human sciences are different, and could give the impression to a casual observer that there can be no shared meaning of what measurement is in such diverse fields. However, a closer look reveals a common condition – the possibility of explicit justification – and thus helps clarify how the “reasonableness” criterion of the VIM's definition of measurement is applicable across the sciences.

References

- [1] Tal E 2015 Measurement in science *The Stanford Encyclopedia of Philosophy* Zalta E N (ed.)
- [2] Hansson S O 2017 Science and Pseudo-Science *The Stanford Encyclopedia of Philosophy* Zalta E N (ed.)
- [3] Duncan O D 1984 *Notes on Social Measurement – Historical and Critical* (New York: Russell Sage Foundation)
- [4] Michell J 1999 *Measurement in Psychology: A Critical History of a Methodological Concept* (Cambridge: Cambridge University Press)
- [5] JCGM 2012 *JCGM 200:2012 International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM)* 3rd ed (2008 version with minor corrections) (Sèvres: Joint Committee for Guides in Metrology) (www.bipm.org/en/publications/guides/vim.html)
- [6] Mari L, Maul A, Torres Irribarra D and Wilson M 2017 Quantities, quantification, and the necessary and sufficient conditions for measurement *Measurement* **100** 115-121
- [7] Bentley J P 2005 *Principles of Measurement Systems* (Harlow: Pearson)
- [8] Doebelin E 2003 *Measurement Systems: Application and Design* (New York: McGraw-Hill)
- [9] Narens L 2012 *Theories of Meaningfulness* (New York: Routledge)
- [10] Markus K and Borsboom D 2013 *Frontiers of Validity Theory: Measurement, Causation, and Meaning* (New York: Routledge)
- [11] Mari L 2013 A quest for the definition of measurement *Measurement* **46** 2889-95
- [12] Mari L 2003 Epistemology of measurement *Measurement* **34** 17-30
- [13] Mari L, Carbone P, Giordani A and Petri D 2017 A structural interpretation of measurement and some related epistemological issues *Stud. Hist. Philos. Sci.* **65-66** 46-56
- [14] Mari L, Carbone P and Petri D 2012 Measurement fundamentals: A pragmatic view *IEEE Trans. Instr. Meas.* **61** 2107–2115

- [15] Morawski R Z 2013 An application-oriented mathematical meta-model of measurement *Measurement* **46** 3753-3765
- [16] Rossi G B and Crenna F 2018 A formal theory of the measurement system *Measurement* **116** 644-651
- [17] Petri D, Mari L and Carbone P 2015 A structured methodology for measurement development *IEEE Trans. Instr. Meas.* **64** 2367-2379