preprint of a paper published in:
Measurement: Interdisciplinary Research and Perspectives, 15, 2, 73-75, 2017, doi:/10.1080/15366367.2017.1367246

Can formal methods provide (necessary and) sufficient conditions for measurement?

Luca Mari

While starting from one, experimentally documented, observation – validation methods might produce positive outcomes even in cases in which their behavior should be the opposite – this thought-provoking paper by Andy Maul arises several important foundational issues, that he aptly introduces and discusses.

My initial reaction after having read the paper has been: yes, of course; the instruction “We are interested in your intuitions – that is, your gut feelings – so please respond to each item based on your first reaction. There are no right or wrong answers.” deliberately lowered the acceptance threshold of the responders, who were thus explicitly solicited to provide their answers even in semantically critical cases (hypothetical counterexample: what would the outcomes have been with an alternative instruction such as “Think twice before responding” and with the possibility of choosing a “don’t know” option?).

And my second reaction has been then to ponder: could such an anomalous case happen also in measuring a physical property / quantity? This question has not only a subjective reason – my work mainly stems from physical measurement, sometimes called “metrology” – but is also justified by the acknowledgement that metrology has a tradition that makes it a reference, though surely not in terms of simple imitation, for social measurement (by the way, I will follow here the terminology of the International Vocabulary of Metrology (VIM) [JCGM 2012]).

Hence: could this happen in metrology? No, I suppose, because in a metrological case the interaction between the object under measurement and the measuring system does not involve any semantic-dependent relationship. Of course, also in that case the measurer needs to understand the measurement procedure, including the instructions on how to make the interaction happen in the expected way, but this does not relate to the content of the interaction, i.e., to the signal that flows through the measuring chain.

Vice versa, measurement setups in (at least cases like this in) psychometrics are such that the correct understanding of the content of the items by the responder, who is the subject / object under measurement, is a substantial precondition of measurement. Were measurement uncertainty taken into account (in metrology this is more and more perceived as a critical condition: no (at least implicit) uncertainty? no measurement), any doubt on the partial understanding, or misunderstanding, of the responder could be modeled in terms of an influence quantity whose presence would affect the overall measurement uncertainty, and therefore, inversely, the quality of the measurement result, and therefore its significance.

Left without a semantic support in the questions (Study 1) on gavagai, or (Study 2) stated as nonsensical sentences, or (Study 3) with completely missing data, but asked to respond nevertheless, the subject plausibly found her / himself in a situation that could be described as a regression to pure syntax, where there is literally nothing to understand and the (syntactic) information is entirely in the identification of differences, à la Shannon [Shannon, Weaver 1949]. In this perspective both the responders’ behavior and the (semantically and pragmatically) wrong positive outcomes of the validation methods obtain an explanation: the responders adopted a syntactically consistent behavior (until the input pattern does not change, the same response is produced, whatever it is), and the validation methods fell in the trap of treating syntactic consistency as semantic validity.

If this simple analysis is correct, Andy’s experiment epitomizes the deficiencies of purely formal, input - output methods of measurement characterization, modeling, and validation. They can provide necessary conditions to differentiate measurement from, say, guessing, but hardly such conditions can be assumed to be also sufficient. And, of course, necessity without sufficiency may well generate the situation of tests that pass though they should fail.

If the measurement process remains within a black box, and everything is known of it are (its preconditions and) its outputs, operationalism cannot be avoided (remember the radical conclusion by Dingle, who defined measurement as “any precisely specified operation that yields a number” [Dingle 1950]). It is my opinion, first articulated in [Mari 2000] and then developed, e.g., in [Mari et al 2012], that the so-called representational theories of measurement [Krantz et al 1971] are affected by the same issue. No need to acknowledge that representation and uniqueness theorems are... theorems, indeed, but their sufficiency to

characterize an experimental setup as a measurement is controversial, to say the least, but in the case ‘measurement’ is defined as a generic morphic mapping (a move with an operational flavor, isn’t it?). In other terms, such theories are, unproblematically, theories of morphic representation, but they are not sufficient to be specifically theories of measurement. I suppose indeed that the individuals involved in Andy’s questionnaire could justify their responses exactly in terms of a consistent, morphic behavior, but this is not sufficient to validate the process as a measurement of anything.

Worse, even necessity is not a yes-or-no condition here, as highlighted by the fact that any threshold in the decision process (like, e.g., the significance level in hypothesis testing set to 0.05) has a conventional component: the concept ‘bad measurement’ – where the low quality might manifest in an only partial fulfillment of the morphic conditions for the given scale – is not oxymoronic. A bad measurement is still a measurement.

This “weak necessity which is not also sufficiency” signals that “an account of what it means for the attribute to vary (i.e., what meaning can be attached to claims about there being “more” or “less” of it, between and possibly within individuals)” (I am quoting Andy) is a key condition for measurement. Indeed, the information provided by measurement is not only about (syntactic) differences, but, according to Bateson, “differences that make a difference” [Bateson 1972].

Let me conclude with a quotation from the Guide to the expression of uncertainty in measurement (GUM): “Although this Guide provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value.” [JCGM 2012]. It is written about techniques of uncertainty evaluation, but it applies – I think – to all formal methods whenever they are tools for measurement.

References

G. Bateson, Steps to an ecology of mind, University of Chicago Press, 1972.

H. Dingle, A theory of measurement, The British Journal for the Philosophy of Science, 1, 5–26, 1950.

JCGM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM, originally published in 1993), Joint Committee for Guides in Metrology, 2008 <http://www.bipm.org/en/publications/guides/gum.html>.

JGCM 200:2012, International Vocabulary of Metrology (VIM) – Basic and General Concepts and Associated Terms (2008 edition with minor corrections), Joint Committee for Guides in Metrology, 2012 <http://www.bipm.org/en/publications/guides/vim.html>.

D. H. Krantz, R. D. Luce, P. Suppes, A. Tversky, Foundations of measurement, Vol. 1, 2, 3, Academic Press, 1971, 1989, 1990.

L. Mari, Beyond the representational viewpoint: a new formalization of measurement, Measurement, 27, 2, 71–84, 2000.

L. Mari, P. Carbone, D. Petri, Measurement fundamentals: a pragmatic view, IEEE Trans. Instr. Meas., 61, 8, 2107–2115, 2012

C. E. Shannon, W. Weaver, The mathematical theory of communication, University of Illinois Press, 1949.