Measurement 133 (2019) 41–46

Contents lists available at ScienceDirect

Measurement

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

On the meaning of measurement uncertainty

Fabien Grégis [a,b,*]

[a] The Cohn Institute for the History and Philosophy of Science and Ideas, Humanities Faculty, Tel Aviv University, Ramat Aviv, Tel Aviv 6997801, Israel
[b] Université Paris Diderot, Sorbonne Paris Cité, laboratoire SPHERE UMR 7219, Bâtiment Condorcet, case 7093, 5 rue Thomas Mann, 75205 Paris Cedex 13, France

ARTICLE INFO

Article history:
Received 2 September 2018
Received in revised form 26 September 2018
Accepted 28 September 2018
Available online 29 September 2018

Keywords:
Measurement uncertainty
Vocabulary of metrology
Epistemology of measurement

ABSTRACT

This article discusses the definitions of "measurement uncertainty" given in the three editions of the *International Vocabulary of Metrology* (VIM) and a fourth definition which was suggested for the next edition of this document. It is argued that none of the definitions is satisfying. First, a thorough definition of measurement uncertainty should supply an explanation about the meaning of the concept, which is missing from the VIM2&3. Secondly, when provided, the meanings are not accurate enough: the VIM1 version is flawed and the possible future definition appeals to an overly restrictive notion of "belief". Alternative options are then proposed, based on the conclusions that (i) measurement uncertainty is a statement – an inference – made by the experimenter about the measurand; (ii) it is only based on what is accessible to one's knowledge, and therefore rests on the hypothesis that no unknown systematic error affects the measurement.

1. Introduction

Measurement uncertainty is a key concept of metrology and scientific practice. Since the end of the 1970s and the recommendation INC-1 [1], lots of efforts have been invested in metrology in order to standardize the practice of uncertainty analysis and to reach an international consensus about the quantitative determination of uncertainties. This process led to the publication of the "GUM family of documents" (term borrowed from [2]), which comprises a series of guides, the *Guide to the expression of uncertainty in measurement* (GUM) [3] and its supplements [4–7], in close relation with the *International Vocabulary of Metrology* (VIM) [8–10], another metrology guide setting up the definitions of critical terms of measurement.

These recent developments were instrumental in clarifying a number of frequently debated issues, among which the difference between measurement error and measurement uncertainty [11]. However, judging by the liveliness of the debates regarding measurement uncertainty today, whether inside or outside the metrological community, it appears that the adequate definition of the concept remains challenging. Despite documents like the GUM and the VIM, a lot of users (including science students [12,13]) may struggle to grasp the meaning of the concept and to appreciate its importance in experimental science. This article intends to reconsider the definitions that were laid in the successive editions of the VIM. I argue that they are either incomplete or partly inaccurate, and I suggest an alternative orientation for the revision of these definitions.

The definition of measurement uncertainty in the *International Vocabulary of Metrology* (VIM) has changed significantly since the first edition in 1984. The successive definitions are listed in Table 1 below, including a possible new definition suggested by the JCGM working group in charge of the conception of a future fourth edition of the document.[1] None of these definitions is satisfying. First, the VIM2 and VIM3 definitions are incomplete. They merely point to what uncertainty purports to describe – a dispersion of values – while I claim that a thorough definition of measurement uncertainty should supply an explanation about the meaning of the concept. Secondly, when the definitions do indeed try to delineate the meaning of measurement uncertainty, as is the case in the VIM1 and the VIM4(?), the meanings provided are not accurate enough. The next sections will be first devoted to demonstrate these two shortcomings. I will then suggest a direction for improvement.

Section 1 gives a general overview of the problem by comparing the definitions provided by the different editions of the VIM. Section 2 argues that the VIM2 and VIM3 definitions are incomplete because they avoid discussing the meaning of measurement uncertainty. Section 3 argues that the VIM1 definition is flawed, and that it cannot easily be improved by acknowledging the probabilistic

* Address: The Cohn Institute for the History and Philosophy of Science and Ideas, Humanities Faculty, Tel Aviv University, Ramat Aviv, Tel Aviv 6997801, Israel.
E-mail address: fabien.gregis@etu.univ-paris-diderot.fr

[1] Because of the provisional status of the latter definition, I will thereafter designate it as the "VIM4(?)" definition, as is done in [14].

https://doi.org/10.1016/j.measurement.2018.09.073
0263-2241/

Table 1
Definitions of "measurement uncertainty" in the three editions of the *International Vocabulary of Measurement* (VIM) [8,9,10] and a possible new definition for the future [14].

VIM1 (1984)	Uncertainty of measurement: an estimate characterizing the range of values within which the true value of a measurand lies
VIM2 (1993)	Uncertainty of measurement: parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could be reasonably attributed to the measurand
VIM3 (2008)	Measurement uncertainty: non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used
VIM4(?)	Measurement uncertainty: parameter (or parameters) characterizing how well the (essentially unique) true value of the measurand is believed to be known

nature of measurement uncertainty. Section 4 argues that the definition suggested for the next edition of the VIM, while making a sensible point, remains too narrow. Finally, in Section 5, I make a proposition for improving the definitions criticized previously.

2. VIM2 and VIM3 definitions avoid dealing with the meaning of measurement uncertainty

The VIM2 and VIM3 consider measurement as a process of *attribution* of values to the measurand [10, p. 16]. Their definitions of measurement uncertainty focus on descriptive features that are relevant in the process of attribution. The main difference between the two definitions resides in that the former uses conditional ("values that *could* be attributed") while the latter uses infinitive ("values *being* attributed"). This is because measurement uncertainty is not considered by the VIM2 as part of the result itself. In the VIM2, the measurement result is the central value (the "value attributed to a measurand, obtained by measurement" [9, p. 23]), and measurement uncertainty is then "associated with the result" in order to characterize it [9, p. 23]. On the contrary, as explained in [14, p. S150], the VIM3 definition makes measurement uncertainty an integral part of the measurement result: since a "measurement result" is a "set of quantity values being attributed to a measurand (...)" [10, p. 19], and since measurement uncertainty is the "dispersion" of this set of values, then measurement uncertainty is simply a descriptive parameter of the measurement result.

If we leave out this difference for the moment, it appears that both definitions are arguably appropriate regarding what they intent to focus on. They have the merit to clarify that measurement uncertainty is not a *property* of the measurand or of the measurement. Indeed, it is a *statement*, made by the experimenters *about* the result, with regard to their theoretical and experimental knowledge. In that sense, measurement uncertainty differs in nature from "measurement error" (or "measurement accuracy") which are (unknown) objective features of the measurement result[2] (respectively the measurement process), and are independent of the knowledge of the experimenters. The VIM2&3 succeed in capturing this aspect of measurement uncertainty (contrarily to the VIM1 definition). The VIM3 statement could very well remain at the basis of a revised definition for measurement uncertainty.

However, the VIM2&3 definitions leave the users without any clue about the meaning of the concept. None of them discloses how the attributed values are to be interpreted. As explained by Charles Ehrlich [14], they do not even tell us *what* one is uncertain *about*.

This predicament might very well be intentional on the part of metrologists, in order to avoid controversy on the meaning of the concept. The VIM2&3 definitions somehow play a role of placeholders: they merely state the existence of a concept, whose meaning is yet to be stipulated. Such a choice allows for the existence of different frameworks for the analysis and interpretation of measurement uncertainty. This could be fitting considering the context in which the VIM2&3 were composed: intense debates have emerged within the community of metrologists during these past decades with regard to the nature of the probabilities and the statistical methods to be used in the context of measurement uncertainty evaluation [15].

Indeed, the traditional commitment to a frequentist interpretation of probabilities – where probabilities are long-term relative frequencies of occurrence of events – has been seriously challenged since the beginning of the 1970s. Many articles of the contemporary literature in metrology advocate an alternative interpretation of probability, often called "epistemic", in which probabilities describes rational degrees of belief.[3] This movement is especially visible in the GUM and its first supplement. Epistemic probabilities were first introduced as an answer to a specific shortcoming of the traditional frequentist approach, namely its inability to provide a probabilistic treatment of systematic errors. An increasing number of metrologists, statisticians and practitioners now advocate a fully Bayesian account of measurement as the best framework for the conceptualization of measurement uncertainty – which is part of the rationale for the introduction of the VIM4(?) definition [14].

Each of these frameworks, frequentist and Bayesian, suggests a different way to formalize, calculate, and interpret measurement uncertainty. This makes the latter's definition even more challenging. The direction taken in the VIM2&3 may thus be understood as a neutral standpoint avoiding to take side in this controversy, being indifferent to the formalism. This position might seem reasonable; but in that case, the VIM2&3 simply do not define the concept of measurement uncertainty.

Such a *status quo* is arguably unsatisfactory and, as a matter of fact, the VIM2&3 definitions contrast with both VIM1 and VIM4(?) definitions, which do tell something about the meaning of measurement uncertainty by introducing the so-called "true value" of the measurand.[4] I claim, however, that these definitions are not fully satisfying either. I will start by arguing that the VIM1 definition is incorrect.

3. The VIM1 definition of measurement uncertainty is flawed

The VIM1 defines measurement uncertainty as "an estimate characterizing the range of values within which the true value of a measurand lies". This definition cannot work: it is deterministic,[5] and admittedly, knowledge is uncertain. A concrete example of why the definition fails is found in precision physics. A closer look at the "adjustments of the physical constants" (see for example [20]) offers to track the recommended values of the physical constants since 1929. The past and present recommended values of the Planck con-

[2] According to [14], the JCGM even tends to consider now that "measurement error should be defined as a quantity, which itself has a value" (p.S152).

[3] See [16] for a philosophical introduction to the concept and [17] for an explanation of its application to metrology. The GUM devotes an appendix to this question, see [3], p.57. Klaus Weise and Wolfgang Wöger's paper [18] is often referred to as one of the seminal articles introducing Bayesian methods in measurement uncertainty analysis.

[4] As [14, p.S150] explains, "plain and simple, the GUM tells us that what we are uncertain about is the value, meaning true value, of the measurand." The concept of "true value" of a quantity is without a doubt tricky and controversial. However, I will take Ehrlich's statement as a sound starting point and will not question its validity here. I will purposely avoid any debate about the concept of "true value", which I discuss further in [19].

[5] One might argue that the definition only makes measurement uncertainty an "estimate" here. However, this specification is ambiguous and the rest of the definition still makes measurement uncertainty a characterization of where the true value effectively lies.