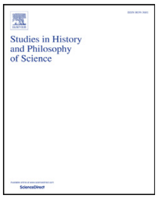
ARTICLE IN PRESS

Studies in History and Philosophy of Science xxx (2017) 1–12

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

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



The evaluation of measurement uncertainties and its epistemological ramifications

Nadine de Courtenay*, Fabien Grégis

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 3 February 2016
Received in revised form
31 May 2016
Available online xxx

Keywords:
Measurement
Uncertainty
Error
Bayesian statistics
Frequentist statistics
Social epistemology

ABSTRACT

The way metrologists conceive of measurement has undergone a major shift in the last two decades. This shift can in great part be traced to a change in the statistical methods used to deal with the expression of measurement results, and, more particularly, with the calculation of measurement uncertainties. Indeed, as we show, the incapacity of the frequentist approach to the calculus of uncertainty to deal with systematic errors has prompted the replacement of the customary frequentist methods by fully Bayesian procedures. The epistemological ramifications of the Bayesian approach merge with a deep empiricist mood tantamount to an “epistemic turn”: measurement results are analysed in terms of degrees of belief, and central concepts such as error and accuracy are called into question. We challenge the perspective entailed by this epistemic turn: we insist on the centrality of the concepts of error and accuracy by underlining the intentional character of measurement that is intimately linked to the process of correction of experimental data. We further circumvent the difficulties posed by the classical analysis of measurement by stressing the social rather than the epistemic dimension of measurement activities.

1. Introduction

Measurement science has been in a state of ferment in the past two decades. Catalysed by the scientific and technical advances of the last century, and by the requirements of economic globalization, it has experienced a period of clarification and reform. Two important guides have been published in order to harmonize the vocabulary, concepts, and measurement practices of metrology (the science of measurement) at the international level – the *International Vocabulary of Metrology* (*VIM*) and the *Guide to the Expression of Uncertainty in Measurement* (*GUM*) –, and a deep revision of the international system of units, the SI, is underway.

Part of the reason for this recent activity is the revamping of the statistical methods used to deal with experimental data, and, more particularly, with the calculation of measurement uncertainties. How to calculate uncertainty has been a major subject of discussions in metrology at least since the middle of the twentieth century. The *GUM*, published in 1993, sought to resolve these discussions by providing probabilistic bases to the calculus, but the proposal was not found satisfactory. It did, however, generate a number of lively debates, which prompted today’s profound transformation of the analysis of measurement data by replacing the classical frequentist methods with Bayesian approaches. The epistemological ramifications of the Bayesian approach merge with a deep empiricist mood pervading the metrological community to instigate a far-reaching revision of the way metrologists conceive of measurement. This revision is tantamount to an “epistemic turn”: measurement results, the traditional touchstones of scientific objectivity, are analysed in terms of degrees of belief, and central normative concepts such as error and accuracy are called into question.

After taking stock of the way measurement error and measurement uncertainty are introduced in the analysis of experimental data, we will explain how the transformation of the metrological conception of measurement originates in the attempt to provide a probabilistic treatment of systematic errors which are of paramount concern in measurement issues. Indeed, the epistemic interpretation of uncertainty, and of measurement as a whole, is designed to avoid the conundrum posed by the classical account of measurement when it claims to assess the correctness of a measurement result by comparing it with the unknown, and unknowable true value of the target quantity. The determination to elude entities that cannot be given empirically, such as the true value of a quantity, results in dismissing the notion of error and

---

* Corresponding author.
E-mail addresses: nadine.decourtenay@univ-paris-diderot.fr (N. de Courtenay), fabien.gregis@etu.univ-paris-diderot.fr (F. Grégis).

http://dx.doi.org/10.1016/j.shpsa.2017.05.003

Please cite this article in press as: de Courtenay, N., & Grégis, F., The evaluation of measurement uncertainties and its epistemological ramifications, Studies in History and Philosophy of Science (2017), http://dx.doi.org/10.1016/j.shpsa.2017.05.003

replacing the requirement for accuracy with that of a rational expression of our knowledge. We will challenge the downgrading of error and accuracy and propose an analysis that stresses the pragmatic and social, rather than the epistemic dimension of measurement. Our approach will suggest that the difficulties attached to the objective evaluation of the quality of a measurement result, and therefore to the concepts of error and accuracy, can be circumvented when one thoroughly takes into account the intentional character of measurement, and acknowledges that the expression of a measurement result involves the positing of a true value as a regulative ideal guiding an activity of correction involving the interactive criticism of a community of agents with a common target. It becomes then possible to conceive of accuracy in a new way; not as the impossible static appreciation of the closeness of the result to a true value, but as a feature related to the reliability of a process of correction anchored in the objectives, values and norms embedded in the social framework underlying measurement activities.

2. Analysis of the variability of measurement indications: measurement errors and measurement uncertainty

2.1. The singular nature of measurement data

A measurement datum is a singular entity. It is the result of a concrete interaction between a physical system bearing the quantity one wants to measure (the length of a particular end gauge, the velocity of light in the vacuum) and a particular experimental setup, in a particular environment, at a particular time, according to a particular procedure. The information derived from such an interaction on the quantity of interest is inevitably entwined with information pertaining to the setup, the environment, and the procedure followed. The question immediately arises of how this datum can be used to give adequate information on the quantity of interest when the quantity is set in a different experimental environment, under different circumstances. How can one obtain from such a measurement datum, information that is valid outside of the particular context in which the datum was produced? Here, one is confronted at the most basic level with the question of how to transform indications of a singular and local nature into measurement results that convey a general, public knowledge of quantities that can be meaningfully and reliably shared.

The first condition the information should meet is that of *communicability*. The chief precondition of communicability is the unit of measurement. As Giordani and Mari (2011) have pointed out, measurement is an experimental process by which a concrete, empirically given quantity *Q*, known by acquaintance, which cannot always be shared, gets expressed by a quantity value *{Q}[Q]*, where *{Q}* is a number and *[Q]* a measurement unit, and thus turns out to be known also by description. Knowledge about the quantity can thus be communicated to distant operators. This description is accomplished by assigning the quantity to a class, identified by *{Q}*, within a classification determined by the publicly defined unit *[Q]*. The assignment is achieved by experimentally comparing the concrete quantity with a standard materializing the unit.[1]

Our main concern, in this paper, will be with the other condition that the information must fulfil in order to be valid beyond the context of its production: a quantity value obtained in a given set of circumstances should be *comparable* with a quantity value of the same concrete quantity obtained in different circumstances; in other words, it should be projectable outside of the experimental context in which it was produced in order to be able to be compared with other evaluations of the same quantity obtained in different circumstances, with theoretical predictions or with technical specifications.

As already mentioned, rough indications obtained in a particular experiment do not satisfy this condition. Their singular nature, the fact that they are tied to a particular context, shows up in their variability: provided one operates with instruments of sufficient resolution, a measurement process will yield different indications when it is repeated. This variability is a straightforward obstacle to comparability; it can be analysed and rectified, but never entirely: it is not possible to completely do away with the context of production. We will see that the precondition that makes it possible to deal with the remaining variability, and allows the handling of comparisons by giving the means to make judgements of sameness and difference, is the "uncertainty" associated with the measurement result and derived from the analysis of variability. In order to perform its function, the uncertainty must be quantified. As a consequence, public, usable measurement results should always be stated with their associated uncertainty.

2.2. From measurement errors to measurement uncertainty

The variability of measurement indications manifests itself in two very different ways. It appears, firstly, when a series of repeated measurements of the same physical system in identical conditions (one says, in "conditions of repeatability") is realized: if the resolution of the experimental set-up is good enough, the measurement indications obtained in these successive experiments will not be the same: they will show a dispersion. Another kind of variability appears when one undertakes to measure the same quantity in distinct experiments, differing either in the measurement principle applied, or in the instruments involved, in the environment or other circumstances (one talks then of "conditions of reproducibility"). Contrary to what happens in the first case, this kind of variability is not observed within the context of a single experiment; it only shows up when one confronts the indications gathered from a variety of different experiments.

The classical way to handle the problem posed by the variability of measurement data is to postulate the uniqueness of measurement results and, by so doing, to introduce the notion of error of measurement.[2] The rationale for such a postulate lies in "our vague and general […] theory of physical objects" which leads us to think that physical properties, and therefore the quantities measured, do not change in conditions of repeatability (Kyburg, 1992, p. 77). We then *explain* the difference between the indications y_i (gathered in the case of a direct measurement) and the unique true target value *TV* of the quantity intended to be measured, called the "measurand",[3] by resorting to the concept of measurement error. The error e_i bearing on the indication y_i is then given by (1) $y_i = TV + e_i$. The true value of the measurand and the measurement errors introduced are thus theoretical concepts suggested by our expectations and theories, which play a normative role in orienting our analysis of measurement experiments. The theoretical status of errors of measurement appears even more salient when one considers the incompatibility between the indications obtained in conditions of reproducibility; one then draws on the hypothesis that quantities are conserved or can be reproduced in different places at different times, and can be determined by resorting to different laws and measurement principles.

---

[1] For more on these issues, see Giordani and Mari (2011).

[2] For a comprehensive account of the notion of measurement error, see Boumans and Hon (2014).

[3] Joint Committee for Guides in Metrology (2012, p. 17).

Please cite this article in press as: de Courtenay, N., & Grégis, F., The evaluation of measurement uncertainties and its epistemological ramifications, Studies in History and Philosophy of Science (2017), http://dx.doi.org/10.1016/j.shpsa.2017.05.003