Modeling measurement: error and uncertainty

Luca Mari [1 *] *and Alessandro Giordani* [2]

[1] Università Cattaneo - LIUC, Castellanza (VA), Italy
[2] Università Cattolica, Milano, Italy

Abstract
In the last few decades the role played by models and modeling activities has become a central topic in the scientific enterprise. In particular, it has been highlighted both that the development of models constitutes a crucial step for understanding the world and that the developed models operate as mediators between theories and the world. Such perspective is exploited here to cope with the issue as to whether error-based and uncertainty-based modeling of measurement are incompatible, and thus alternative with one another, as sometimes claimed nowadays. The crucial problem is whether assuming this standpoint implies definitely renouncing to maintain a role for truth and the related concepts, particularly accuracy, in measurement. It is argued here that the well known objections against true values in measurement, which would lead to refuse the concept of accuracy as non-operational, or to maintain it as only qualitative, derive from a not clear distinction between three distinct processes: the metrological characterization of measuring systems, their calibration, and finally measurement. Under the hypotheses that (1) the concept of true value is related to the model of a measurement process, (2) the concept of uncertainty is related to the connection between such model and the world, and (3) accuracy is a property of measuring systems (and not of measurement results) and uncertainty is a property of measurement results (and not of measuring systems), not only the compatibility but actually the conjoint need of error-based and uncertainty-based modeling emerges.

Keywords: *measurement; models in science; error; uncertainty; measurand.*

1 Introduction
The relation between error and uncertainty in measurement is a complex subject, in which philosophical assumptions and modeling methods are variously intertwined. Both the concepts ‘error’ and ‘uncertainty’ aim at keeping into account the experimental evidence that even the best measurements are not able to convey definitive and complete information on the quantity intended to be measured. The pragmatic approach prevailing in the literature on instrumentation and measurement often leaves today the distinction between error and uncertainty implicit, as it were just a lexical issue in which a single, stable concept is differently expressed.[1] While at the operative level this interchangeability can find some justifications, e.g., the rule customarily adopted for the analytical propagation of errors and uncertainties is indeed the same, whether such concepts are alternative or complementary is a particularly important topic, also because of the claimed current “change in the treatment” of the subject, “from an Error Approach (sometimes called Traditional Approach) to an Uncertainty Approach” (quoted from the *International Vocabulary of Metrology*, 3rd edition, “VIM3” henceforth (JCGM 2012)).
In this paper, after having introduced the problem and presented the general background of our analysis in the context of the realist vs. instrumentalist opposition in philosophy of science (Section 2), we propose a solution by exploiting the concept of models as mediators between the informational level of the propositions and the ontological level of the world (Section 3). In particular, it is shown how the concept of truth can be suitably applied to propositions characterizing the relation between a quantity value and a *quantity construed as an entity specified in the context of given model*, while the concept of (un)certainty suitably applies to characterize the relation between such a modeled quantity and a *quantity as an empirical entity specified by reference to a concrete object*. This is the basis for an in-depth analysis of the complex connection between truth and (un)certainty and the development of a framework in which these concepts are combined.
After a brief revision of the operational contribution given by the *Guide to the expression of uncertainty in measurement* (“GUM” henceforth (JCGM 2008a)) (Section 4), a simple model of the operation performed by a measuring instrument is proposed, and then compared with the two other basic processes in which the instrument is involved: calibration and metrological characterization (Section 5). The analysis of the three processes highlights that a concept of (operative) true value, and therefore of error, can be introduced, by pushing the

* 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.

1 As an example, consider the following quotation from a well known textbook (Taylor 1997): “All measurements, however careful and scientific, are subject to some uncertainties. Error analysis is the study and evaluation of these uncertainties [...]. The analysis of uncertainties, or “errors,” is a vital part of any scientific experiment”. “Error” and “uncertainty” seem to be used here interchangeably.

position that measurement is able to convey "(operative) pure data". By relaxing some of the given hypotheses, it is finally shown (Section 6) that in a broader and more realistic context measurement could be understood as a knowledge-based process, where both (operative) truth and (un)certainty play a significant role.

2 Two perspectives on error and uncertainty

The relation between error and uncertainty in measurement is so deeply rooted in foundational issues that even the basic related terminology is somehow controversial and has to be preliminarily agreed upon. Still, even a trivial case, such as the measurement of the length of an usual object as a table, suits the purpose of exemplifying the required concepts. We will call:

– the table a, the *object under measurement* (it could be a phenomenon, an event, a process, ...);
– the length Q, a *general quantity* (sometimes called attribute, observable, property, kind of quantity, undetermined magnitude, ...);
– the length of the table q_a, *an individual quantity* (sometimes called magnitude), which is here the quantity intended to be measured, the *measurand* for short;[2]
– an entity v such as 2.34 m, a *quantity value*.

The basic objective of measurement m can be then described as the assignment of the best (in a sense to be specified) quantity value to a given measurand, i.e., to a general quantity considered of the object under measurement.

$$q \xrightarrow{m} v$$

The theoretical assumptions concerning the measurand are crucial for the relation between error and uncertainty in measurement. In particular, in a hypothetical spectrum of theoretical options about the knowledge of the quantitative characters of the world, two extremes can be identified: a *realist* standpoint, according to which knowledge depends on how the world is structured in itself, and an *instrumentalist* standpoint, according to which knowledge depends on how the world is determined by our conceptual structures. This alternative significantly affects the interpretation of the nature of measurement and the involved entities.

According to the realist standpoint:[3]	According to the instrumentalist standpoint:[4]
both general and individual quantities exist independently of measurement;	*neither* general nor individual quantities exist independently of measurement;
individual quantities are *actually* related by numerical ratios;	individual quantities can be *operationally* related by numerical ratios;
once an individual quantity is *selected* as a unit, all other individual quantities of the same general quantity	once an individual quantity is *introduced* as a unit, all other individual quantities of the same general quantity

2 Whether the measurable entities are quantities or, more generically, properties is an issue for measurement science, but it is immaterial here. Hence we will maintain the traditional, restricted, version in our discussion, which may be generalized without difficulty to the case of ordinal and classificatory (or "nominal") entities. It should also be noted that individual quantities are instances of general quantities (e.g., the length if this table *is a* length). This justifies the usual loose terminological habit of leaving the distinction implicit.

3 The realist standpoint rests on a tradition, dating back to the Pythagorean philosophers and championed by scientists as Galilei and Kepler, which conceives the world as having a definite mathematical structure: the world "is written in mathematical characters" (Galilei) since "numbers are in the world" (Kepler).

4 The instrumentalist standpoint rests on a tradition, dating back to the Protagorean philosophers and championed by philosophers as Carnap, which conceives the world as being known in virtue of the adoption of a conceptual framework: "We introduce the numerical concept of weight by setting up a procedure for measuring it. It is *we* who assign numbers to nature. The phenomena themselves exhibit only qualities we observe. Everything numerical [...] is brought in by ourselves when we devise procedures for measurement" (Carnap 1966, Part II, p.100). In addition, Carnap construes quantitative concepts as essential tools for introducing quantitative laws and views such laws as essential tools for making previsions: "The most important advantage of the quantitative law [...] is not its brevity, but rather the use that can be made of it. Once we have the law in numerical form, we can employ that powerful part of deductive logic we call mathematics and, in that way, make predictions". This conception of the role of quantitative concepts and laws justifies the use of the term "instrumentalism" with respect to the proposed standpoint. It is also worth noting that such a standpoint is based on operationalist assumptions (Bridgman 1927) about the definition of scientific concepts and that is typically adopted by scholars influenced by the view that that "pure data does not exist" (Kuhn 1970), because "data are always theory laden" (Hanson 1958).

are *determined* by a number;	can be *represented* by a number;
measurement is a process aimed at *discovering* the measurand, where the quantity value states the result of such a discover.	measurement is a process aimed at *assigning* a quantity value to the measurand.

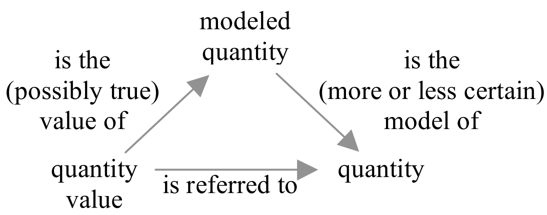
Let us suppose that an experimental situation is given where repeated measurements, which are assumed to be performed on the same or similar objects and with respect to the same quantity under specified conditions, provide different quantity values. The two standpoints interpret the situation as follows.
According to the realist standpoint, since quantities exist in the world, an object a is characterized with respect to a general quantity Q by a definite individual quantity q_a, to which a true value $v(q_a)$ is in principle associated. The fact that repeated measurements provide different values for q_a is justified in terms of measurement errors, which make knowing the true value impossible *in practice*. The assumption of the existence of the true value $v(q_a)$ leads to the idea that such value can be better and better approximated as the measurement process is enhanced and the effects of measurement errors are reduced.
According to the instrumentalist standpoint, since quantities do not exist in the world, a is not characterized with respect to Q by any definite individual quantity q_a. The fact that repeated measurements provide different values for the measurand is accounted for in terms of an uncertainty which makes knowing true values impossible *in practice and in principle*. The assumption of the non-existence of the true value $v(q_a)$ leads to the idea that measurement uncertainty is an irreducible feature of measurement and characterizes "the dispersion of the quantity values being attributed to the measurand, based on the information used" (JCGM 2012).
In the context of measurement science, the concept of *true quantity value* is an unavoidable element in the first perspective, as witnessed by its foundational role in the classical theory of errors, and an element of no significance in the second one.
This tension has led to a critical attitude towards the concept of true quantity value, which is now used, if even outside papers and books it is used at all, always stressing that it is an ideal, unknowable element. With this acknowledgment, the VIM3 has proposed the above mentioned idea of two "approaches", the "Error Approach" and the "Uncertainty Approach", contrasted with each other with respect to the expected objective of measurement: in the Error Approach, to determine an estimate of the true value of the measurand which is as close as possible to that unique true value; in the Uncertainty Approach, to decide an interval of values which can justifiably be assigned to the measurand. Hence, this opposition is further interpreted as related to the alternative: unique true quantity values vs. intervals of quantity values, where sometimes, possibly for maintaining a reference to the tradition, the hybrid concept 'intervals of true quantity values' is adopted.[5] The change in the article – *the* true value vs. *a* true value – might be taken as a lexical symbol of this conundrum.
All these elements support the hypothesis that error and uncertainty are incompatible concepts, which cannot be reconciled even at the operative level, so that assuming the latter implies definitely renouncing to maintain a role for truth, and therefore error, in measurement. Is this actually the case? Our claim here is that this incompatibility of "approaches" is inconsistent and unjustified, and that it can and should be overcome.

3 Framing the perspectives

The realist *vs.* instrumentalist opposition on the role of error and uncertainty in measurement can be inscribed in a broader contraposition concerning knowledge as such,[6] which can be roughly presented as follows. The realist standpoint emphasizes that the conceptual framework we use to describe the world can be abstracted from the observable world and provides us with information about the structure of the world itself. The instrumentalist standpoint, on the other hand, emphasizes that such conceptual framework is constructed by ourselves and then interpreted on the world in such a way that the interpretation determines the reference of the concepts. These perspectives are both, although differently, related to the problem of the truth. According to the realist standpoint, which assumes the building blocks of the propositions to be obtained from the world, any proposition has a definite *truth value*, depending on its actual correspondence with a considered "portion of the world". On the contrary, the instrumentalist standpoint considers that the building blocks of the propositions are introduced by us, so that a proposition has only an *aptness value*, expressing its capability of allowing us to successfully interact

5 On this respect the "Uncertainty Approach" is rather delicate, since the contemporary presence of different truths on the same subject might appear a violation of the principle of non-contradiction. This topic will not further discussed here.
6 See (Ladyman 2007) for a general introduction to the debate.

with the world.
Some recent developments in philosophy of science[7] have suggested a way to conciliate these perspectives by focusing on the role of *models* in the knowledge of the world, and on the function of measurement, a crucial knowledge tool, accordingly.[8] Indeed, in the last decades the view of models as mediators for understanding and interacting with the world has been variously stressed: the current picture, while too complex and controversial to be accounted for here, can be outlined in its general traits as follows:[9]
– any portion of the world of interest is conceptualized by means of a model, in which the pertinent objects and properties characterizing that portion as a system under consideration are isolated;
– this model is used both as a theoretical tool for interpreting our concepts and as an operational tool for studying the corresponding portion of the world.
This way of conceiving the relation between models and portions of the world permits us to dismiss both (i) the position according to which propositions allowing for a successful interaction with the world are true with respect to the world and (ii) the position according to which propositions are never true but only more or less successful. The proposal here is that a proposition can be true with respect to a model and a model is more or less similar to the world, and as such it is more or less successful. Indeed, let us assume that a portion of the world can be conceptualized by means of a model and that a model can be described by means of propositions. In addition, let us assume that
– to be successful is a property that can be significantly ascribed only to instruments, and particularly to models, which are non-linguistic entities,
whereas
– to be true is a property that can be significantly ascribed only to propositions, i.e., to specific linguistic entities.
Then a proposition is neither more or less successful nor true with respect to the world, but it can be a true description of a successful model, and thus true with respect to that very model.
By interpreting measurement in terms of models the concepts of true quantity value and measurement uncertainty can be combined without being forced to interpret uncertainty in terms of doubt about the truth of a measured value. The basic idea is that truth refers to the attribution of a quantity value to a modeled quantity, while uncertainty characterizes both such attribution and the degree of similarity between the model of that quantity and the actual, empirical quantity, and therefore also to their combination, i.e., the attribution of a quantity value to the quantity.



Hence, the attribution of a quantity value to a quantity involves uncertainty both because the quantity is modeled in such a way that it is difficult, or even impossible, in principle to know its value (e.g., because it is modeled as its value is in a set of real numbers) and because the model of the quantity is not necessarily accurate (e.g., because influence quantities are not properly taken into account).
A simple, but paradigmatic, example comes from the way in which the problem of measuring the area of a paper sheet is solved. A twofold idealization is customarily introduced, according to which the object under measurement is modeled as a rectangle and the measurand is defined as a general quantity, area, of that object and that can be evaluated by positive real numbers and which is not influenced by other quantities. The solution is then obtained by measuring the length of the two sides of the sheet and taking their product. Hence, in this model the object under measurement has an individual area, whose true value is obtained by multiplying the lengths of its sides. Still, since such a model is an idealization of the actual sheet, which is surely not "perfectly rectangular", to convey a correct information about the measurand it is necessary to estimate the possible discrepancy between

7 See, in particular, (Giere 1988, ch. 3), (Giere 2010), and (Morgan 1999, ch. 2).
8 This role has been already acknowledged in (Kuhn 1961) and (Suppes 1962).
9 The introduction of models as mediators between the conceptualization of an entity and the entity itself allows reproducing the opposition realism vs. instrumentalist at the level of the theoretical entities whose existence is required in the construction of the model. For an extensive discussion of this issue in the recent epistemological debate see (Psillos 1999), where scientific realism is defended, and (van Fraassen 1989), where a strong version of the opposite position, constructive empiricism, is developed.

the value obtained in virtue of the ideal measuring procedure and the value, indeed the set of values, obtained in virtue of the actual measuring procedure, and then to express this discrepancy in terms of uncertainty.
This conclusion shows that the realist and the instrumentalist standpoints are not required to be thought of as fundamentally opposed. Still, this position, despite its interest, constitutes only a partial solution to the problem of the relations between error and uncertainty in measurement, and this for two reasons: first, because the recent developments of the operational perspective in metrology have introduced new motivations for casting doubts about the concept of a unique true value; second, because the analysis can be deepened to show that in the very case of measurement the concept of true value has application with reference not only to modeled quantities, but also to actual quantities, i.e., to the world.

4 Focusing on the operational perspective

A landmark in the development of a widespread adoption of uncertainty modeling in measurement has been the publication, in 1993, of the GUM, aimed at establishing "general rules for evaluating and expressing uncertainty" (JCGM 2008a), building on a recommendation issued in 1980 by a Working Group on the Statement of Uncertainties convened by the Bureau International des Poids et Mesures (BIPM) and approved in 1981 by the Comité International des Poids et Mesures (CIPM), the world's highest authority in metrology. To this purpose a concept of measurement uncertainty was to be preliminarily adopted: innovating the tradition, the GUM focuses on operational, instead of epistemological or ontological, issues, and stresses in particular:
(1) the construal of the concept of measurement uncertainty;
(2) the introduction of the concept of definitional (there called "intrinsic") uncertainty;
(3) the elimination of the concept of true quantity value as identical with the concept of quantity value.
About the point (1), measurement uncertainty is interpreted as uncertainty on the quantity value to be attributed to the measurand, where indeed the GUM recommends a measurement result to be a pair (measured quantity value, measurement uncertainty), from which the usual interval form, $v \pm \Delta v$, can be obtained.[10] Hence, measurement uncertainty becomes a parameter "characterizing the dispersion of the quantity values being attributed to a measurand" (JCGM 2012), instead of a parameter estimating the distance of the quantity value from the true one.

> The definition of uncertainty of measurement [...] is an operational one that focuses on the measurement result and its evaluated uncertainty. However, it is not inconsistent with other concepts of uncertainty of measurement, such as (i) a measure of the possible error in the estimated value of the measurand as provided by the result of a measurement; (ii) an estimate characterizing the range of values within which the true value of a measurand lies. [...] Although these two traditional concepts are valid as ideals, they focus on unknowable quantities: the "error" of the result of a measurement and the "true value" of the measurand (in contrast to its estimated value), respectively. Nevertheless, whichever concept of uncertainty is adopted, an uncertainty component is always evaluated using the same data and related information.[11]

The classical distinction between random and systematic errors – which attributes being random or systematic as a feature of error itself – is set aside, and a new classification "Type A" vs. "Type B" is adopted, which, most importantly, refers not to the sources of uncertainty, but "to the way in which numerical values are estimated", either "by statistical methods" or "by other means" respectively. This switch is claimed to give an operative solution to a long standing problem: "random and systematic errors [...] have to be treated differently [and] no rule can be derived on how they combine to form the total error of any given measurement result". Hence, the concept of error can be avoided here and the philosophical assumptions concerning truth and true values discharged: there is no necessity to make a reference to a supposed true value, since there is no possibility to evaluate the distance between the estimated value and the true value. Rather, measurement uncertainty is acknowledged as deriving from several possible sources (and the GUM lists several of them: a) incomplete definition of the measurand; b) imperfect realization of the definition of the measurand; c) nonrepresentative sampling – the sample measured may not represent the defined measurand; d) inadequate knowledge of the effects of environmental conditions on the measurement or imperfect measurement of environmental conditions; e) personal bias in reading analogue instruments; f) finite instrument resolution or discrimination threshold; g) inexact values of measurement standards and reference materials; h) inexact values of constants and other parameters obtained from external

10 The Supplement 1 of the GUM (JCGM 2008b) acknowledges that this hypothesis can be generalized, and deals with measurement results in the form of probability distributions.
11 See (JCGM 2008a, 2.2.4). According to the first of such "other concepts", uncertainty would be "a measure of error", a not so clear interpretation (e.g., is error considered here a non-quantitative concept, quantitatively expressed as / by means of uncertainty?). The subject will not further discussed here.

sources and used in the data-reduction algorithm; i) approximations and assumptions incorporated in the measurement method and procedure; j) variations in repeated observations of the measurand under apparently identical conditions) and nevertheless all of them are to be formalized as standard uncertainties, i.e., standard deviations. This assumption leads to a single rule for combining all components of uncertainty – the so called law of propagation of uncertainty – which is thus applicable to components obtained by both Type A and Type B evaluations.

About the point (2), the fact that the information obtained by measurement is supposed to be related to a measurand introduces an unavoidable interpretive component in the measurement problem, which then requires the measurand to be specified. This implies giving a description of both the object under measurement and the environment expected when the measurement takes place. Hence, in principle a measurand cannot be completely described without an infinite amount of information. The consequence is that the measurand, now defined as the quantity *intended* to be measured (JCGM 2012, 2.3), is to be distinguished from the quantity actually subject to measurement, and a definitional uncertainty[12] remains as for this distinction.

About the point (3), the GUM states that the word "true" in "true value" is "unnecessary", because "the «true» value of a quantity is simply the value of the quantity".

These three points highlight some significant tensions in the GUM approach. The concept of definitional uncertainty is introduced, but then deprived of any operative import, since the definitional uncertainty is simply assumed to be negligible with respect to the other components of measurement uncertainty. Under this hypothesis, the value attributed to the measurand can be assumed as "essentially unique" (and note: "essentially", not, e.g., "practically"), so that the expressions "value of the measurand" and "true value of the measurand" are taken as equivalent. On the other hand, precisely this position, where the equation between the "true value" of the measurand and "the best estimate of the value" of the measurand is assumed, leads to the problem of building a conceptually consistent framework. It is surprising to declare both that the true value is eliminable, since unknowable, and identical with the best estimate of the value, which is evidently known. Alternatively, it is surprising to declare both that the true value is the value of the measurand, identical with the best estimate of the value of the measurand, and that this value is not truly representing the measurand. Thus, stating that "true" is redundant seems to be here just a lexical position, which leaves unmodified the problem about true values.

In addition, there is a further price to be paid for this emphasis on the operational side of the problem. The assumed unknowability of true values transfers to another pivotal concept for measurement, the one of measurement accuracy, defined as "closeness of agreement between a measured quantity value and a true quantity value of a measurand" (JCGM 2012, 2.13). As a consequence, then, "the concept 'measurement accuracy' is not a quantity and is not given a numerical quantity value", plausibly an elliptical way to state that were measurement accuracy considered a quantity then its value would be unknowable in its turn because of the reference in its definition to a true value. Still, accuracy is customarily listed among the features of measuring instruments, and a numerical value for it is indicated. The way outs which are sometimes adopted to solve the puzzle, e.g., redefining the concept relating not to a true value but to a "conventional true" value (is then truth assumed to be conventional?) or a (generic) reference value (ISO 1998), do not really help to cope with the general issue.

As a synthesis, while the introduction of the definitional uncertainty, which leads us to improve our initial picture, where a model true value is introduced, can be viewed as a positive contribution of the GUM, the apparently deflationist strategy[13] underlying the elimination of the concept of true value, instead of clarifying the frame, seems to obscure important characters of the measurement process and to neglect the fact that not all quantity values are unknowable, as, e.g., nominal quantity values – which represent design constraints – are known by specification.

5 Rethinking the concepts and their relations

With the aim of further exploring the concepts of measurement error and measurement uncertainty let us introduce an admittedly simplified model of the empirical core of a measurement process, *a measuring instrument*, interpreted as a basic device enabling quantity representation. The VIM3, which defines it a "device used for making measurements, alone or in conjunction with one or more supplementary devices", notes that "a measuring

12 The concept was introduced as "intrinsic uncertainty" in the GUM (and, as mentioned above, the first source of uncertainty listed by the GUM is indeed the "incomplete definition of the measurand"), then re-termed "definitional uncertainty" by the VIM3 (JCGM 2012, 2.27 Note 3).

13 As in the standard example: it is true that "the snow is white" if and only if the snow is white (David 2009), where "the snow is white" and "it is true that the snow is white" convey exactly the same information, and all usages of "true" are of this kind, so that any reference to truth can be eliminated (Stoljar 2010).

instrument may be an indicating measuring instrument or a material measure” (JCGM 2012, 3.1). An indicating measuring instrument (material measures can be omitted in the following discussion) is a “measuring instrument providing an output signal carrying information about the value of the quantity being measured” (JCGM 2012, 3.3), such an output signal being called the indication (JCGM 2012, 4.1). An indicating measuring instrument operates in fact as a generic transducer, which dynamically produces an output quantity q'_{out} as the effect of its interaction with (an object characterized in particular by) an input quantity q_{in}.
A simple example of transducer which can be operated as an indicating measuring instrument is a spring: in response to the application of a force (the input quantity q_{in}), the spring stretches and a length is thus obtained (the output quantity q'_{out}).[14] What does it make a spring a measuring instrument? Three basic conditions.
MC1: the output quantity q'_{out} is assumed to reliably provide information on the quantity being measured q_{in}.
This is a condition of predictable input-output behavior, requiring that (i) the mapping $q_{in} \rightarrow q'_{out}$ can be formalized as a function τ_q by an underlying theory, or at least a black box causal modeling, of such behavior, and that (ii) the transducer which implements τ_q is properly constructed and operated according to a given procedure. The function τ_q corresponds to the *empirical component* of measurement.

$$q_{in} \xrightarrow{\tau_q} q'_{out}$$

MC2: the output quantity value v' corresponding to q'_{out} is assumed to be known.
This condition prevents a never-ending recursive process: were the output quantity subject to measurement in its turn, a further transducer would then be required (a measuring instrument whose input quantity is length in the case of a spring), and for its output quantity this condition should apply. Hence, it is required that (i) the mapping $q'_{out} \rightarrow v'$ can be formalized as a function d_{out}, and that (ii) such function is known. The function d_{out} corresponds to the *evidential component* of measurement.

$$\begin{tikzcd} q_{in} \arrow[r, "\tau_q"] & q'_{out} \arrow[d, "d_{out}"] \\ & v' \end{tikzcd}$$

MC3: an input quantity value v corresponding to the output quantity value v' can be obtained.
This results from the hypotheses that (i) the function τ_q between quantities is mirrored by an invertible function τ_v between quantity values, and that (ii) such function is known. The application of τ_v^{-1} constitutes the core of the *inferential component* of measurement.

$$\begin{tikzcd} q_{in} \arrow[r, "\tau_q"] & q'_{out} \arrow[d, "d_{out}"] \\ v & v' \arrow[l, "\tau_v^{-1}"'] \end{tikzcd}$$

Whenever conditions MC1-MC3 are satisfied measurement can be performed, as formalized as a function m obtained by the composition of the empirical, the evidential, and the inferential components, $m = \tau_v^{-1} \bullet d_{out} \bullet \tau_q$.

$$\begin{tikzcd} q_{in} \arrow[r, "\tau_q"] \arrow[d, dotted, "m"'] & q'_{out} \arrow[d, "d_{out}"] \\ v & v' \arrow[l, "\tau_v^{-1}"'] \end{tikzcd}$$

As a synthesis, this model of the measurement process, even though extremely simple, shows that several components have to be coordinated to establish the relation between a measurand and a measurand value.

14 As the example suggests, the individual quantities q_{in} and q'_{out} are generally of different kinds, i.e., they are instances of different general quantities, Q and Q', where *in* is the object under measurement and *out* is usually the transducer itself, which changes its state as the result of its interaction with the object *in*.

5.1 Indications and indication values (discussion on MC2)

MC2 *is crucial for the present discussion*. Let us analyze it a little bit more thoroughly. First of all, the relation between quantities and quantity values is clearly a fundamental topic for measurement, which is indeed aimed at associating a measured quantity value to a quantity, the measurand.[15] On the other hand, there are some clues that such relation is still open for discussion and better clarification.[16]

Back to transducers, and therefore to measuring instruments: are indications then quantities or quantity values? If the lexical confusion is not justified, in this case the conceptual superposition is acceptable, precisely because of MC2 (in the following we will remove the ambiguity by using "indication" for a quantity and "indication value" for a quantity value). In any measuring instrument there must be a point in which the relation between a quantity and a quantity value is assumed as given, as a *pure datum* in an operative sense, i.e., as an entity that, although possibly further specifiable, can be recognized and recorded without any additional interpretive or inferential process. According to MC2 this point is where indication values are produced, thus justifying the term "evidential" introduced above for this component. In fact, measuring instruments are designed so to make the mapping from indications to indication values straightforward, being typically implemented as a process of pattern recognition, performed by human beings or technological devices: the observation of coincidence of marks, the classification of an electric quantity to a quantized level to which a digital code is associated, the numbering of right answers of a test, and so on.[17] All of them are assumed as unproblematic processes leading to values for the given quantity, i.e., the instrument indication, under the assumption that a pattern recognition of sufficiently low specificity is both truthful and its outcome cannot be further refined in the context. In this way, the information conveyed by the mapping from indications to indication values is the best one which can be achieved by means of the instrument. Accordingly, the obtained indication value can be properly said the true quantity value of the corresponding indication. Of course, this is a revisable, operative truth, so that the term *operative true quantity value* could be adopted to denote this entity, which has to be systematically distinguished from the *model true quantity value* introduced before. On the other hand, such operative true values are indication values, not measurand ones (length values instead of force values in the case of a spring). Hence, such truth is still not sufficient for measurement.

5.2 Transduction and calibration (discussion on MC3)

Since the measurement function m is obtained as $\tau_v^{-1} \bullet d_{out} \bullet \tau_q$, a critical problem to ensure the correctness of m concerns the way in which the mapping τ_v^{-1} from indication values to measurand values is obtained. Such problem is solved by means of instrument *calibration*. Interestingly, the description of the instrument operation for measurement and calibration is the same: the instrument interacts with an object, and an indication is obtained as the result of the transduction of an input quantity of the object. Still in a simplified model, the conditions for calibration are as follows.

CC1: a set of reference objects ref_j is available, each of them providing an input quantity $q_{ref,\,j}$ of the transducer and thus producing a corresponding output quantity $q'_{out,\,j}$.

$$q_{ref,\,j} \xrightarrow{\tau_q} q'_{out,\,j}$$

CC2: the quantity value v_j for each of such reference quantities $q_{ref,\,j}$ is assumed to be known.

---

15 Surprisingly, there seems to be some confusion even on this topic. For example, (Bentley 2005, p.3) states: "The input to the measurement system is the true value of the variable; the system output is the measured value of the variable. In an ideal measurement system, the measured value would be equal to the true value." The assumption that the input of a measurement system is a quantity value, instead of a quantity, appears to be an obvious mistake.

16 For example, (JCGM 2009) states that "in future editions of [the VIM] it is intended to make a clear distinction between the use of the term «error» as a quantity and as a quantity value. The same statement applies to the term «indication». In the current document such a distinction is made. [On the other hand, the VIM3] does not distinguish explicitly between these uses". The reference is plausibly to some inconsistencies in the VIM3, as when it is stated that "indications, corrections and influence quantities can be input quantities in a measurement model", thus assuming indications as quantities, against the definition of 'indication' as "quantity value provided by a measuring instrument or a measuring system".

17 "In carrying out measurements there is a tendency to reduce the immediate sensory observations, which of course can never be eliminated, to the safest and most exact among them, namely spatio-temporal coincidences (in particular, one tries to do without the subjective comparison of colors and light intensities). Any mensuration should ultimately ascertain, so one wishes, whether a mark on one scale (a movable pointer or such) coincides with a certain mark on another scale" (Weyl 1949, p.145).

$$\begin{tikzcd} q_{ref,j} \arrow[r, "\tau_q"] \arrow[d, "d_{ref}"'] & q'_{out,j} \\ v_j & \end{tikzcd}$$

Also in this case a never-ending recursive process must be prevented: were the reference quantity subject to measurement in its turn, a measuring instrument would be required, and the problem should arise again with respect to that instrument. Hence, it is required that (i) the mapping $q_{ref,j} \rightarrow v_j$ can be formalized as a function d_{ref}, and that (ii) such function is known. The function d_{ref} corresponds to the *evidential component* of calibration.
Together with MC2, CC1 and CC2 ensure that information on both $q_{ref,j}$ and $q'_{out,j}$ is provided.

$$\begin{tikzcd} q_{ref,j} \arrow[r, "\tau_q"] \arrow[d, "d_{ref}"'] & q'_{out,j} \arrow[d, "d_{out}"] \\ v_j & v'_j \end{tikzcd}$$

Calibration is then aimed at producing the map $v'_j = \tau_v(v_j)$, as a set of pairs, $\{\langle v_1, v'_1\rangle, \langle v_2, v'_2\rangle, ...\}$, or an analytical synthesis of them, for example resulting from the hypothesis of linearity of τ_v, so that $v'_j = \alpha + \beta v_j$ for given parameters α and β.

$$\begin{tikzcd} q_{ref,j} \arrow[r, "\tau_q"] \arrow[d, "d_{ref}"'] & q'_{out,j} \arrow[d, "d_{out}"] \\ v_j \arrow[r, dotted, "\tau_v"] & v'_j \end{tikzcd}$$

This highlights the fundamental role of calibration for the inferential part of measurement, which is based on the relation encoded in τ_v. In its turn, the construction of τ_v is based on the evidence concerning the available operative *pure data*, i.e., about indication values, as given by d_{out}, and about reference values, as given by d_{ref}.

5.3 References and reference values (discussion of CC2)

Hence, CC2 *is the second crucial element in this discussion*. Like MC2, it supposes the operative availability of unproblematic values, assigned by convention, or through a chain of responsibility delegation, typically guaranteed by a calibration hierarchy (sometimes called a traceability chain) from a primary measurement standard. Together with MC2, it ensures the possibility of mirroring the empirical mapping τ_q, which connects quantities, i.e., empirical entities, with the informational mapping τ_v, which connects quantity values, i.e., informational entities. It is in this way that in a calibrated measuring instrument *the empirical component is reliably linked to the evaluation component*.

This point is delicate. In some presentations the whole problem of measurement is introduced by assuming that the input of the measuring system is the true measurand *value* (see, e.g., the quotation from (Bentley 2005) in a previous footnote), that experimental errors "hide" in the mapping to indication values, with the consequence that error theory / analysis has the purpose of estimating such input value despite the superposed errors. This is a overly simplified, and actually misleading, position: being empirical devices, measuring instruments interact with (quantities of) objects, not values. In particular, a value for the measurand is the final outcome of the process, not its starting point (but under the above mentioned realist hypothesis that "numbers are in the world"). Any criticism to true values based on this assumption is thus well founded but, as we are going to argue, this does not imply that truth, and therefore error, must vanish from the scope of measurement.

5.4 Measurement

As just considered, in both calibration and measurement the empirical mapping τ_q is operated. The difference is about the interpretation and the purpose of this operation:

– in calibration the mapping τ_v is unknown and is looked for on the basis of the knowledge of the values v_j associated to references and of indication values v'_j;

– in measurement the input quantity value, v, is unknown and is looked for on the basis of the knowledge of τ_v and of the indication value v' obtained by transducing q_{in}.

Measurement exploits calibration information by inverting it:

$v = \tau_v^{-1}(v')$

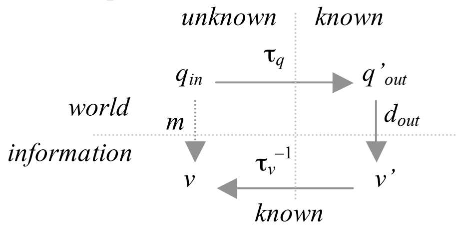
Hence, the measurement m of the measurand q_{in} is performed by:[18]

1. empirically transducing q_{in} to an indication $q'_{out} = \tau_q(q_{in})$;
2. mapping q'_{out} to the corresponding value $v' = d_{out}(q'_{out})$;
3. mapping v' to the value $v = \tau_v^{-1}(v')$ attributed to the measurand.

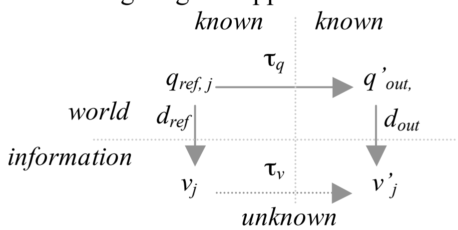
In synthesis:

$v = m(q_{in}) = \tau_v^{-1}(d_{out}(\tau_q(q_{in})))$

According to this simplified model, the whole picture of *measurement* is as follows.



The condition on the basis of which measurement is possible is the construction of τ_v, as achieved in the process of *instrument calibration*, to which the following diagram applies.

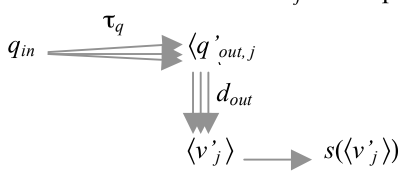


From a foundational point of view, the significance of these pictures is apparent.

5.5 Metrological characterization

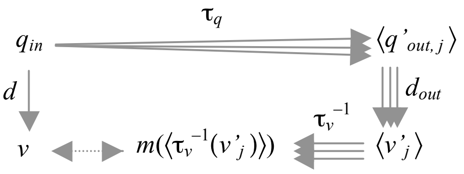
Together with calibration and measurement, the empirical mapping τ_q is exploited also in a third kind of process: the metrological characterization of measuring instruments (the VIM3 calls "verification", "provision of objective evidence that a given item fulfills specified requirements" (JCGM 2012, 2.44), a similar process). Indeed, neither calibration nor measurement are aimed at producing information on the metrological capability of a measuring instrument. To this goal two basic processes (in particular) can be designed:

- an input quantity q_{in} is repeatedly applied to the measuring instrument and a scale statistic (e.g., sample standard deviation, $s(.)$) on the sample of the produced indication values v'_j is computed:



- an input quantity q_{in} whose value v is assumed to be known is repeatedly applied to the calibrated measuring instrument, a location statistic (e.g., sample mean, $m(.)$) on the sample of the produced measured quantity values $\tau_v^{-1}(v'_j)$ is computed, and then v and the value of such statistic are compared:

18 Apparently this model does not comply with the customary condition that measurement implies comparison of quantities. In fact, in some cases measurement is based on the synchronous comparison of the object under measurement and a measurement standard, paradigmatically as performed by a two-pan balance. The simplified model presented here refers to the operatively much more frequent situation in which the object under measurement and the measurement standard are not required to be simultaneously present, under the hypothesis that the measuring instrument is stable enough to maintain in measurement the behavior which was characterized in calibration. In this sense, the model is about measurement performed as asynchronous comparison of the object under measurement and the measurement standard.



The first process does not require q_{in} to be provided by a measurement standard nor the measuring instrument to be calibrated, but is critically based on the stability of q_{in}. Its aim is to convey information on the stability of the mapping τ_q, a property usually called *measurement precision* (according to the VIM3 the "closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions" (JCGM 2012, 2.15)). Under these assumptions, any non-null value for the given scale statistic has to be considered as the indicator of *errors* in the transduction behavior of the measuring instrument.

The second process is more demanding, since it requires not only the stability of q_{in} but also the knowledge of its value v, as typically obtained by means of a measurement standard, together with the calibration of the measuring instrument. Its aim is to convey information on the stability of the mapping τ_v, and therefore of the calibration itself, a property called *measurement trueness* (according to the VIM3 the "closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value" (JCGM 2012, 2.14)). Under these assumptions, any difference between v and the location statistic has to be considered as the indicator of *errors* introduced by the fact that the calibration information is not correct anymore.[19]

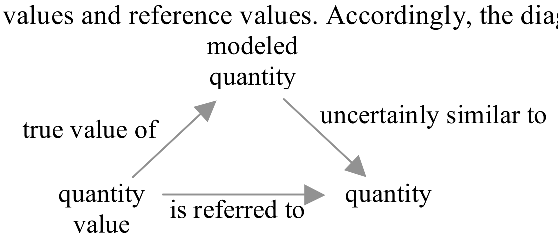
It is fundamental to note here that:

– the mentioned errors refer to a measuring instrument but are identified and evaluated in a process which is not a measurement; only under a (customarily reasonable, in fact) hypothesis of stability of the measuring instrument, this information about its erratic behavior can be assumed to hold also for measurement;

– precision and trueness, and then accuracy, are features related to the measuring instrument behavior, and more generally to measurement if the measurement procedure is taken into account, and not to measurement results, although the information they convey is appropriately exploited in the evaluation of measurement data to assign a measurement result.

6 Conclusions: reshaping the framework

The realist and the instrumentalist perspectives can be composed, at a first glance, by allowing for model true values and interpreting uncertainty as representing the acknowledged discrepancy or dissimilarity between a model and the modeled portion of the world. Still, we have also highlighted that (1) it is possible to be uncertain as to the model to choose in order to analyze a given portion of the world, and that (2) it is possible to admit true values with respect to indication values and reference values. Accordingly, the diagram:



does not provide a sufficiently general account of the significance of the concepts of true value and uncertainty in measurement: some uncertainty affects the modeled quantity, and the concept of truth can be applied to values of

19 The distinction in the VIM3 between trueness and accuracy, defined as "closeness of agreement between a measured quantity value and a true quantity value of a measurand" (JCGM 2012, 2.13), is not completely clear, given that true quantity values are specific kinds of reference quantity values. In a different context (ISO 1998), building on measurement precision and trueness, measurement accuracy is thought of as an overall indicator, which summarizes the information conveyed by both precision and trueness. How such information can be synthesized is outside the scope of this paper.

some actual quantities too.
As a conclusion, the following synthesis, asking for further developments, can be offered.
– A measurand can be modeled at different levels of specificity (in the sense according to which the measurand ‘length of this metal rod at the temperature of 30 °C’ is more specific than ‘length of this metal rod’). The chosen level of specificity is typically not the most specific one, so that there is an uncertainty as to which model, among the ones that specify the chosen model, would be the most similar one.
– A connection between quantities and quantity values is the outcome of the measurement of measurands but is also assumed, as a definitional assignment, in the cases of indication quantities and reference quantities. Since such quantities are produced by a designed process, it appears to be legitimate to consider the values assigned to them as their operative true values.
In addition, if the simplified model *of* measurement introduced so far is embedded in a more realistic context, the role of models *in* measurement has to be taken into account. The idea is that the quality of measurement results is affected not only by measurement errors but also by other causes, which in the specific context of the given measurement are not empirically controllable and therefore can be evaluated only on the basis of given interpretive hypotheses. The resulting effects are expressed in terms of *uncertainty* in measurement results. Among such other causes there are the following ones (these are just hints: each of them could require a much more thorough analysis).
– *Modeling the calibration hierarchy*. The working standards exploited in instrument calibration customarily are calibrated in their turn, through a calibration hierarchy. While each step of this process is a calibration, the trueness-related errors might not combine linearly. Hence, instead of trying to conceive a complex super-model including the whole calibration hierarchy, the expert knowledge on the resulting effects can be exploited, and elicited in terms of uncertainty of the quantity value of the working standard.
– *Modeling the transduction process*: *identifying the laws*. The transduction behavior of the measuring instrument is not perfectly characterized. If the transducer is assumed to obey a parametric law (typically: it is linear, i.e., the transducer sensitivity is constant in the measuring interval) its calibration is greatly simplified but at the price of an approximation, which can be taken into account in terms of uncertainty of the involved quantity values.
– *Modeling the transduction process*: *closing the system*. The transducer exploited in measurement is not perfectly selective: the transduction process is perturbed by several influence quantities, so that its output depends not only on the stated input quantity. While in principle each influence quantity might be measured in its turn and its effects properly characterized and then eliminated, this process would lead to a never-ending recursive process, due to the fact that in the measurement of an influence quantity some influence quantities should be taken into account. Typically, some simplifying hypotheses are assumed on the effects of the influence quantities, expressed as an uncertainty of the indication value.
– *Modeling the measurand*: *identifying the quantity intended to be measured*. The fact has to be admitted that the input quantity of the transducer might not be the measurand, i.e., that the quantity subject to measurement is not the quantity intended to be measured. In such a case a model has to be adopted to infer the measurand value from the available information and, of course, this model could be acknowledged as implying some simplifications, whose effects can be expressed as an uncertainty on the stated measurand value.
– *Modeling the measurand*: *defining the quantity intended to be measured*. The measurand as such might be defined only up to a given approximation, resulting in a definitional uncertainty, “the practical minimum measurement uncertainty achievable in any measurement of a given measurand” according to the VIM3. The metrological model of the measurand is the source of an important kind of uncertainty, which cannot be eliminated by means of experimental means. If it is impossible to define *the* quantity that is intended to be measured, then it is, even in principle, impossible to determine *its* quantity value.
These examples highlight that measurement uncertainty can be assumed as an encompassing concept by means of which the quality of measurement results is expressed by taking into account both the effects of measurement errors and the approximations due to measurement-related models. As a consequence, “the error approach” and “the uncertainty approach” to measurement are not only compatible but actually both required for an appropriate evaluation of measurement data: measurement errors are a component of the usually broader set of causes of measurement uncertainty.

References

J.P. Bentley, Principles of measurement systems, Harlow: Pearson, 2005[4].
P.W. Bridgman, The logic of modern physics, New York: Macmillan, 1927.
R. Carnap, Philosophical foundations of physics, New York: Basic Books, 1966.

M. David, The Correspondence Theory of Truth, in: E.N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Fall 2009 Edition) (http://plato.stanford.edu/archives/fall2009/entries/truth-correspondence).
R. Giere, Explaining science: A cognitive approach, Chicago: University of Chicago Press, 1988.
R. Giere, An agent-based conception of models and scientific representation, Synthese, 172, 269-281, 2010.
N. Hanson, Patterns of discovery, Cambridge: Cambridge University Press, 1958.
ISO 5725-1, Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions, International Organization for Standardization, 1998.
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).
JCGM 101:2008, Evaluation of measurement data – Supplement 1 to the “Guide to the expression of uncertainty in measurement” – Propagation of distributions using a Monte Carlo method (http://www.bipm.org/en/publications/guides/gum.html).
JCGM 200:2012, International Vocabulary of Metrology – Basic and general concepts and associated terms (VIM), 3rd Edition (2008 version with minor corrections), Joint Committee for Guides in Metrology, 2012 (http://www.bipm.org/en/publications/guides/vim.html).
JCGM 104:2009, Evaluation of measurement data – An introduction to the ‘Guide to the expression of uncertainty in measurement’ and related documents, Joint Committee for Guides in Metrology, 2009.
T.S. Kuhn, The function of measurement in modern physical sciences, Isis, 52, 161–193, 1961.
T.S. Kuhn, The structure of scientific revolutions, Chicago: University of Chicago Press, 1970.
J. Ladyman, Ontological, epistemological, and methodological positions, in: T. Kuipers (ed.), General philosophy of science: Focal issues, Amsterdam: Elsevier, 2007; ch. 6, pp. 303-376.
M.S. Morgan, M. Morrison (eds.), Models as mediators: Perspectives on natural and social science, Cambridge: Cambridge University Press, 1999.
S. Psillos, Scientific realism: How science tracks the truth, London: Routledge, 1999.
D. Stoljar, N. Damnjanovic, The deflationary theory of truth, in: E.N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Winter 2010 Edition) (http://plato.stanford.edu/archives/win2010/entries/truth-deflationary).
P. Suppes, Models of data, in: E. Nagel, P. Suppes, A. Tarski (eds.), Logic, Methodology, and Philosophy of Science, Stanford: Stanford University Press, 1962; pp. 252-261.
J.R. Taylor, An introduction to error analysis – The study of uncertainties in physical measurements, Sausalito: University Science Books, 1997^2.
B.C. Van Fraassen, Laws and symmetry, Oxford: Oxford University Press, 1989.
H. Weyl, Philosophy of mathematics and natural science, Princeton: Princeton University Press, 1949.