Beyond the representational viewpoint: a new formalization of measurement

Luca Mari
Libero Istituto Universitario C.Cattaneo
C.so Matteotti, 22 – 21053 Castellanza (VA) – Italy
Tel: +39(0)331 5721 – Fax: +39(0)331 572320 – Email: lmari@liuc.it

Abstract: the paper introduces and formally defines a functional concept of a measuring system, on this basis characterizing the measurement as an evaluation performed by means of a calibrated measuring system. The distinction between exact and uncertain measurement is formalized in terms of the properties of the traceability chain joining the measuring system to the primary standard. The consequence is drawn that uncertain measurements lose the property of relation-preservation, on which the very concept of measurement is founded according to the representational viewpoint. Finally, from the analysis of the inter-relations between calibration and measurement the fundamental reasons of the claimed objectivity and intersubjectivity of measurement are highlighted, a valuable epistemological result to characterize measurement as a particular kind of evaluation.

Keywords: foundations of measurement; measurement theory; calibration and measurement, measurement uncertainty

1. Introduction

In the last decades the concept of measurement has been modeled and formalized according to the so-called *representational viewpoint*. As a particular method of evaluation, measurement has been understood as the homomorphic mapping of an empirical relational structure into a symbolic relational structure. At the basis of the representational viewpoint the hypothesis stands that the condition of being a homomorphism is both necessary and sufficient for an evaluation to be accepted as a measurement. In the words of [1]: «the method of constructing the homomorphism is precisely the measurement procedure».

Therefore the representational viewpoint assumes that:

$$\textbf{evaluations} \supset \textbf{homomorphisms} \equiv \textbf{measurements}$$

The significance of this hypothesis becomes manifest in the process aimed at taking a non-homomorphic evaluation and transforming it into a homomorphism, with the goal to faithfully represent by symbols the empirical information on the attribute under evaluation [2].

Typical examples of such non-homomorphic evaluations are the judgments of preference among alternatives that are intended to conform to an ordinal structure but factually present some intransitive chains, i.e. sequences of the kind $a_1>a_2>\ldots>a_n>a_1$. In these cases the empirical relational structure cannot be homomorphically mapped into any symbolic relational structure characterized by a linear order. A transformation from a non-homomorphic evaluation to a homomorphism leads here to linearize the empirical order by means of a normative intervention in which the evaluator is forced to remove the intransitive preferences. The condition implied in such transformations, and thus in the requirement of homomorphism, is then basically directed to guarantee the consistency of the evaluation.

From an epistemic standpoint, the *necessity* of the condition of homomorphism seems to be well founded, since a non-homomorphic evaluation could be hardly called a measurement. Critical is however the *sufficiency* of such a condition, an issue concerning the very meaning of measurement and its role of activity of knowledge acquisition and expression.

As already remarked, the representational viewpoint is based on the hypothesis that the condition of homomorphism is not only necessary but also sufficient to define the concept of measurement: we claim that the reason of this position is extrinsic, being referred to the fact that those who mainly contributed to the development of such a viewpoint were behavioral scientists. In the context of the social sciences a transformation from a non-homomorphic evaluation to a homomorphism can be indeed very meaningful, particularly from the normative point of view, i.e. to help the subject in setting up an evaluation as a 'correct' measurement. On the other hand, the deemed identity of homomorphic evaluations and measurements leaves some of the specific features of the latter unjustified and however unexplained. «When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge of it is of a meager and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science». Could this well-known statement by Lord Kelvin be referred to *any* homomorphic evaluation? The epistemic importance of measurement, to which Lord Kelvin alludes, derives from its being both *objective and intersubjective*, two characteristics whose origin must be discovered outside the language and its formalisms, directly in the operational component of the measurement itself.

Therefore our basic claim is that:

evaluations ⊃ homomorphisms ⊃ measurements

The purpose of this paper is to discuss and give a sound conceptual basis to this standpoint, and accordingly to propose a formalization of measurement as a *specific* kind of homomorphic evaluation. This formalization will allow us to find the reasons to the claimed objectivity and intersubjectivity of measurement in the adoption of a suitably calibrated measuring system and the reference to a standard set suitably tied to a traceability chain, i.e. more or less directly referred to a primary standard. Furthermore, in this metrological-oriented framework the presence of the uncertainty emerges as inherent to measurement, so that its causes and effects will be discussed and formalized.

The concept of measurement will be introduced at first under the ideal assumption of absence of uncertainty, to make the proposed formalization of the notions of measuring system, calibration, and traceability chain easier to understand. Such concepts will be then extended to the general case of uncertain measurement, thus obtaining an overview of the whole framework here proposed.

2. The representational viewpoint of measurement: its conceptual basics and flaws

«The modern form of measurement theory is representational: numbers assigned to objects/events must represent the relations perceived between the properties of those objects/events» [3]. The formalization usually adopted by the representational viewpoint to express such concepts of 'properties of objects/events' and 'relations between properties' is a set-theoretical one, that generalizes the notions of universal algebra and morphism.

A *relational structure* is an ordered pair $\mathbf{X}=<X,R_X>$ of a domain set X and a set R_X of relations on X. The elements of R_X can have different arities. In particular, the 1-ary relations are subsets of X, and the n-ary operations on X (i.e. the functions $X^n \rightarrow X$) belong to R_X as specific (n+1)-ary relations. Given two relational structures $\mathbf{X}$ and $\mathbf{Y}$, a *homomorphism* from $\mathbf{X}$ to $\mathbf{Y}$ is a mapping $\mathbf{m}=<m,m_R>$, $\mathbf{m}:\mathbf{X}\rightarrow\mathbf{Y}$, where:

* m is a function that maps X into $m(X)\subseteq Y$, i.e. for each element of the domain set there exists at least one, but not necessarily only one, corresponding image element;

* m_R is a function that maps R_X into $m_R(R_X)\subseteq R_Y$ such that $\forall r\in R_X$, r and $m_R(r)$ have the same arity, i.e. for each relation on the domain set there exists one (and it is usually, and often implicitly, assumed: and only one) corresponding image relation,

with the condition that $\forall r\in R_X$, $\forall x_i\in X$, if $r(x_1,\ldots,x_n)$ then $m_R(r)(m(x_1),\ldots,m(x_n))$, i.e. if a relation holds for some elements of the domain set then the image relation must hold for the image elements.

Relational structures can be then thought of as sets-with-constraints, and homomorphisms as mappings among relational structures preserving such constraints.

As an example, let $\mathbf{X}=<X=\{x_1,x_2,x_3\}, R_X=\{r\}>$ being $r=\{(x_1,x_2), (x_2,x_3), (x_1,x_3)\}$ (remember that a n-ary relation on X can be expressed as a subset of the cartesian product set X^n), and $\mathbf{Y}=<Y=\{y_1,y_2,y_3,y_4\}, R_Y=\{s\}>$ being $s=\{(y_1,y_1), (y_2,y_2), (y_3,y_3), (y_4,y_4), (y_1,y_2), (y_2,y_3), (y_3,y_4), (y_1,y_3), (y_1,y_4), (y_2,y_4)\}$. Given the following mappings from $\mathbf{X}$ to $\mathbf{Y}$:

- $\mathbf{m_1}$ such that $m_1(x_1)=y_1$, $m_1(x_2)=y_2$, $m_1(x_3)=y_3$
- $\mathbf{m_2}$ such that $m_2(x_1)=y_1$, $m_2(x_2)=y_1$, $m_2(x_3)=y_1$
- $\mathbf{m_3}$ such that $m_3(x_1)=y_1$, $m_3(x_2)=y_2$, $m_3(x_3)=y_1$

then only $\mathbf{m_1}$ and $\mathbf{m_2}$ are homomorphisms (and the latter is a trivial one), since in the case of $\mathbf{m_3}$ $r(x_2,x_3)$ but not $s(m_3(x_2),m_3(x_3))=s(y_2,y_1)$.

Relational structures and homomorphisms are the basic constructs adopted to formalize the concept of measurement. A relational structure **E** is called *empirical* if its domain set E spans over the thing states under consideration; a relational structure **S** is called *symbolic* if its domain set S spans over a given set of symbols. A homomorphism **m** from an empirical relational structure **E** into a symbolic relational structure **S** is adopted to formalize an attribute taking values in S when evaluated on a thing in E [4].

2.1. An interpretation and a critique of the representational viewpoint

On these bases the measurement procedures are interpreted as a 'method of constructing the homomorphism'. The emphasis in such a construction is on «the qualitative conditions under which a particular representation holds» [1], i.e. on the empirical relational structure and in particular the qualitative relations defined in it (a short remark: 'qualitative' is usually adopted in this case in opposition to 'quantitative', being considered quantitative the relations defined among symbols. Both are doubtful denotations). The key point is the requirement that such empirical relations should be *perceivable*, and factually 'perceived' before measurement. Any measurement would be based on the direct comparison of things, an assumption that is seldom verified in the actual practice, and unnecessary from the theoretical point of view.

This claimed need of direct perception formally corresponds to the requirement that an *extensional definition* is available for the relations in R_E, i.e. $\forall r \in R_E$, where r is n-ary, $\forall e_i \in E$, it is empirically known whether $r(e_1,\ldots,e_n)$ or not. On the other hand, this is not sufficient (and in the following we will suggest that it is neither empirically necessary) to define what should be considered a *meaningful* homomorphism of **E**.

As an example, consider a 'directly perceived' empirical order relation r on the domain $E=\{e_1,e_2,e_3\}$, $r=\{(e_1,e_2),(e_2,e_3),(e_1,e_3)\}$. Let $S=\{1,2,3\}$ be the chosen symbol set. Together with the 'order-preserving' homomorphism $\mathbf{m}:<E,\{r\}>\rightarrow<S,\{<\}>$ such that $m(e_1)=1$, $m(e_2)=2$, $m(e_3)=3$, other homomorphisms can be defined, for example $\mathbf{m}':<E,\{r\}>\rightarrow<S,\{\neq\}>$ such that $m'(e_1)=1$, $m'(e_2)=3$, $m'(e_3)=2$ (the fact that **m**' is actually a homomorphism can be easily verified). The homomorphism **m**' appears to be unusual because there are some couples of elements of E not in relation r but whose images are instead in relation $m'_R(r)$ (for example not

$r(e_2,e_1)$ whereas $m'_R(r)(m'(e_2),m'(e_1))$, since $m'(e_2)=2 \neq m'(e_1)=1$). As a consequence, some properties of r, such as the asymmetry, are not preserved by **m**'.

In coherence to its extensional approach, the representational viewpoint eludes this problem by implicitly specializing the definition of homomorphism and requiring that $\forall r \in R_E$, $\forall e_i \in E$, $m_R(r)(m(e_1),\ldots,m(e_n))$ if *and only if* $r(e_1,\ldots,e_n)$, thus forcing the biconditional implication that is not present in the general definition. It can be easily shown that this condition is equivalent to the intensional requirement that the *same* properties hold for the relations in **E** and **S** (neglecting the presence of the so-called anti-homomorphisms: for example, an empirical strict order can be homomorphically mapped into both '<' and '>'). These 'natural' homomorphisms can be neatly understood from a logical point of view by requiring that both **E** and **S** are *models of the same theory*, here with the role of a measurement scale (for an introduction on the concepts of theory and model, as defined in formal logic, and their application to characterize measurement scales, see [1]).

The homomorphisms satisfying this condition will be called here *model homomorphisms*, m-homomorphisms for short. Therefore, formally:

evaluations ⊃ homomorphisms ⊃ m-homomorphisms

so that our basic claim becomes:

m-homomorphisms ⊃ measurements

It should be clear that our objections to the representational viewpoint are in fact not merely terminological, in reference to the choice of denoting as homomorphisms what are in effect m-homomorphisms, and refer to the strategy assumed to *empirically* characterize the m-homomorphisms within the whole classes of the homomorphisms and evaluations. While the check that $r(e_1,\ldots,e_n)$ implies $m_R(r)(m(e_1),\ldots,m(e_n))$ has a manifest significance (provided that $r(e_1,\ldots,e_n)$ is known, of course), the inverse check that $m_R(r)(m(e_1),\ldots,m(e_n))$ implies $r(e_1,\ldots,e_n)$ seems rather an *ad hoc* technique to guarantee that **m** is a m-homomorphism and not just a homomorphism.

On the contrary, we suggest that the characterization of measurement is *intensional*, being based on the knowledge available about the measurand *before* the accomplishment of the evaluation. Such a knowledge is independent of the availability of any extensional information on the relations in R_E. It is therefore correct that «the method of constructing the homomorphism is precisely the measurement procedure», but the fact is that in the scientific and technological practice the definition of a measurement procedure generally requires to set up a measuring system (MS), and not to directly compare things with each other.

2.2. Evaluation levels

Schematically, three levels of evaluations can be thus envisioned:

L1. the measurand is not identified or however a m-homomorphic comparison among things cannot be performed; these can be called *generic evaluations*;

L2. the measurand is identified but a MS for it is not available; on the other hand a m-homomorphic comparison among things can be performed; these can be called *m-homomorphic evaluations*;

L3. the measurand is identified and a MS for its evaluation is available and adopted; in some cases a m-homomorphic comparison among things could be directly performed, but is more usually obtained in indirect way through the comparison of the measurement results and the back-propagation of the result, so that $r(e_1,\ldots,e_n)$ whenever $m_R(r)(m(e_1),\ldots,m(e_n))$; we claim that these, and only these, are properly *measurements*.

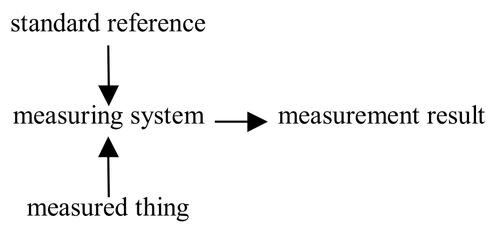
As already discussed, the representational viewpoint seems to be mainly concerned with L2 and the normative conditions leading to the procedural transformation from L1 to L2. To reach L3 a suitable formalization for the concept of a MS must be obtained.

3. The formal concepts of measuring system and measurement

«For the most part, questions about physical measurement are regarded as being in the province of philosophy of physics, not in physics itself […] The construction and the calibration of measuring devices is a major activity, but it lies rather far from the sorts of qualitative theories we examine here» [1]. On the opposite, we deem the theoretical relevance and the unavoidability of such a 'major activity' for the definition of an actual *measurement* theory:

measurement is an evaluation performed by means of a calibrated MS.

An operational concept of a MS calibration must be then introduced and defined. The suggested interpretation of the concept presents it as a sort of counterpart of measurement, as the following diagram sketches:



3.1. Reference relational structures

The formal concept of a MS is rooted on the notion of a reference *standard set* (the term is adopted to generalize the more usual 'standard sequence', to keep into account the algebraically weak scales in which the concept of sequence is not meaningful). A standard set A is a set of thing states conventionally chosen by the operator with the aim of materializing *his knowledge on the measurand*. The cardinality of A is related to the measurand resolution, i.e. the ability to discriminate between thing states with respect to the measurand. A set of relations R_A is defined on A, conveying the information available on the scale in which the attribute is measured. Due to the condition that the elements of A can be discriminated with respect to the measurand, no equivalence relations belong to R_A. Therefore $R_A=\varnothing$ for nominal attributes, R_A={a strict order} for ordinal attributes, and so on.

The presence of one operation *op* in R_A usually allows to constructively generate the standard set A, beginning from a 'reference unit' $\underline{a}$ through a sequence of compositions $2\underline{a}=op(\underline{a},\underline{a})$, $3\underline{a}=op(2\underline{a},\underline{a})$, … This procedure is not allowed if R_A contains only 'pure' relations: in this case (for example for nominal and ordinal scales) the standard set must be given in its extensional form.

Such a basic concept of a standard set must be then specified to keep into account its algebraic structure and the fact that several standard sets are customarily linked with each other to constitute a traceability chain that joins the MS to the primary standard set.

Let $\mathbf{A}_i=<A_i,R_{Ai}>$, $i=1,..,n$ for a given $n\geq1$, be called a i-ary *reference relational structure* (RRS) whenever the whole sequence $\mathbf{A}_1,\ldots,\mathbf{A}_n$ fulfils the following three ideal conditions (note that they will be subsequently relaxed, to keep into account the presence of some uncertainty):

<u>condition 1</u>: for each $i=2,..,n$ the i-ary RRS $\mathbf{A}_i$ is defined in terms of the i–1-ary RRS $\mathbf{A}_{i-1}$, i.e. traced in reference to it, by requiring that there exists a mapping $\boldsymbol{\kappa}_i:\mathbf{A}_{i-1}\rightarrow\mathbf{A}_i$ and that such a mapping is a m-homomorphism. The corresponding function $\kappa_i:A_{i-1}\rightarrow A_i$ is called a 'standard comparison function'. The composition $\boldsymbol{\tau}=\boldsymbol{\kappa}_2\rightarrow\ldots\rightarrow\boldsymbol{\kappa}_n$ expresses the whole *traceability chain*, formalizing the condition that the n-ary RRS $\mathbf{A}_n$ is (indirectly) defined in reference to the primary RRS $\mathbf{A}_1$, and that the resolution of the primary standard set A_1 is not worse than the one of the n-ary standard set A_n, and in particular equal if the function τ is injective;

<u>condition 2</u>: for each $i=1,..,n$ a symbolic relational structure $\mathbf{S}$ is chosen so that for each RRS $\mathbf{A}_i$ a homomorphism $\mathbf{m}_i$, $i=1,..,n$, is defined from $\mathbf{A}_i$ into $\mathbf{S}$ such that m_i is injective, i.e. $\mathbf{m}_i$ is an isomorphism onto its image. This formalizes the possibility to integrally express the empirical information maintained by the RRSs in symbolic terms;

<u>condition 3</u>: a 'thing comparison function' χ is defined from the set E of the thing states under evaluation to the n-ary standard set A_n, $\chi:E\rightarrow A_n$. This formalizes the possibility to adopt $\mathbf{A}_n$ (and indirectly the whole set of the RRSs $\mathbf{A}_i$ belonging to the traceability chain, up to $\mathbf{A}_1$) as a reference in the empirical comparison to the things to be evaluated. Since a direct perception of the relations among things is not required here, χ is not forced to be a homomorphism. In

principle, indeed, no relations must be, and could be, extensionally known on E before measurement.

3.2. A functional standpoint on measuring systems

For any given traceability chain $\boldsymbol{\tau}$, a MS is a system:

* embedding (either empirically or just formally) the RRS $\mathbf{A}_n$;
* calibrated in reference to $\mathbf{A}_1$ by means of $\boldsymbol{\tau}$, as formalized by the homomorphism $\mathbf{m}_n$;
* able to empirically interact with the thing states E, as expressed by the function χ.

The setup and usage of such a MS is then conceptually a (1+3)-step operation (note the denotation of the mappings between empirical states with Greek symbols, and of the mappings involving symbols with Latin symbols):

<u>step 0</u>: *choice of the primary RRS*. An injective homomorphism $\mathbf{m}_1$ is defined from the RRS $\mathbf{A}_1$ into a given symbolic relational structure $\mathbf{S}$ (cf. the condition 2):

$$\mathbf{A}_1 \xrightarrow{\mathbf{m}_1} \mathbf{S}$$

<u>step 1</u>: *calibration*. The homomorphism $\mathbf{m}_2$ is defined by means of the m-homomorphism $\boldsymbol{\kappa}_2$ (cf. the condition 1), by the condition that $\forall a' \in A_2$ such that $a'=\kappa_2(a)$, being $a \in A_1$, $m_2(a')=m_1(a)$:

$$\begin{array}{ccc} \mathbf{A}_1 & \xrightarrow{\mathbf{m}_1} & \mathbf{S} \\ \downarrow{\scriptstyle \boldsymbol{\kappa}_2} & & \uparrow \\ \mathbf{A}_2 & \overset{\mathbf{m}_2}{\cdots\cdots\cdots} & \vdots \end{array}$$

(in the case the function κ_2 is not injective, i.e. if in the transition from the primary to the secondary standard set some resolution is lost, more than one element of A_1 is mapped by κ_2 into the same element of A_2; in this case a suitable aggregation function, for example the average if defined, must be adopted to compute $m_2(a')$ from the set of the values $m_1(a)$).

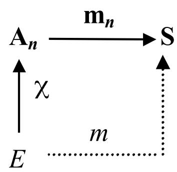
Such an operation is recursively repeated until the RRS $\mathbf{A}_n$ is reached and the homomorphism $\mathbf{m}_n$ is defined:

$$\begin{array}{ccc} \downarrow & & \\ \mathbf{A}_{n-1} & \xrightarrow{\mathbf{m}_{n-1}} & \mathbf{S} \\ \downarrow{\scriptstyle \boldsymbol{\kappa}_n} & & \uparrow \\ \mathbf{A}_n & \overset{\mathbf{m}_n}{\cdots\cdots\cdots} & \vdots \end{array}$$

This homomorphism formalizes the calibration information for the MS, as it is indirectly obtained by the primary standard through the traceability chain;

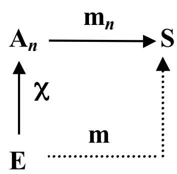
step 2: *empirical interaction with the thing under measurement (data acquisition)*. The RRS $\mathbf{A}_n$ is put in interaction with the thing under measurement in its state $e \in E$. This leads to the selection of a specific element $a \in A_n$, as formalized by the function $a=\chi(e)$ (cf. the condition 3);

step 3: *symbolic selection (data presentation)*. Once calibrated as in step 1 and put in interaction with the thing state as in step 2, the RRS $\mathbf{A}_n$ is adopted to select a symbol in S as the measurand value: $m(e)=m_n(a=\chi(e))$:



where $a=\kappa_n(a')$ for a given $a' \in A_{n-1}$, and $a'=\kappa_{n-1}(a'')$ for a given $a'' \in A_{n-2}$, and $a''= \ldots$ so that $m(e)=m_n(\kappa_n(\kappa_{n-1}(\ldots(\kappa_2(a'''')))))=\tau(a'''')$, being $a'''' \in A_1$ an element of the standard set of the primary RRS.

An equivalence relation eq is induced in E by the function χ such that $eq(e_1,e_2)$ if and only if $\chi(e_1)=\chi(e_2)$, i.e. the evaluated thing states are not distinguished with respect to the measurand. Furthermore, via χ any relation defined on A_n is induced on the quotient set E/eq, and therefore indirectly on the set E itself: for each $r_j \in R_{Ai}$ a corresponding relation r'_j is induced on R_E such that $r'_j(e_1,\ldots,e_n)$ if and only if $r_j(\chi(e_1),\ldots,\chi(e_n))$. This makes $\mathbf{E}=\langle E,\{eq,r'_j\}\rangle$ a relational structure and the mappings χ and m m-homomorphisms $\boldsymbol{\chi}$ and $\mathbf{m}$. Hence:



Provided that a proper identifier **ref** is assigned to the RRS $\mathbf{A}_n$ (for example ‘Richter scale’, or ‘Celsius degrees’, or ‘millimeters’), a measurement result of the thing state e for the attribute **m** is expressed as customarily as ‘$m(e)=s$ **ref** ’ where $s= m_n(\chi(e))$.

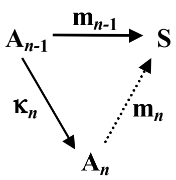
That measurement is formalized as a m-homomorphism is not a constraint on its definition, but only a consequence of the fact that measurement is an operation based on a suitably calibrated MS.

4. Calibration and measurement

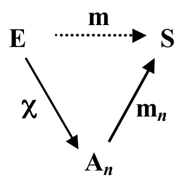
This formalization of a MS highlights several characteristics of the connection between calibration and measurement. Both are operations implying a 3-placed relation: empirical states

× empirical states × symbols. In calibration the known standard states are compared to the unknown MS states to properly assign symbols to the latter. In measurement, the unknown measurand states are compared to the known MS states to properly assign symbols to the former. Therefore calibration and measurement can be thought of as *inverse* operations. Re-expressing the previous diagrams in the usual form of commutative triangles, indeed:

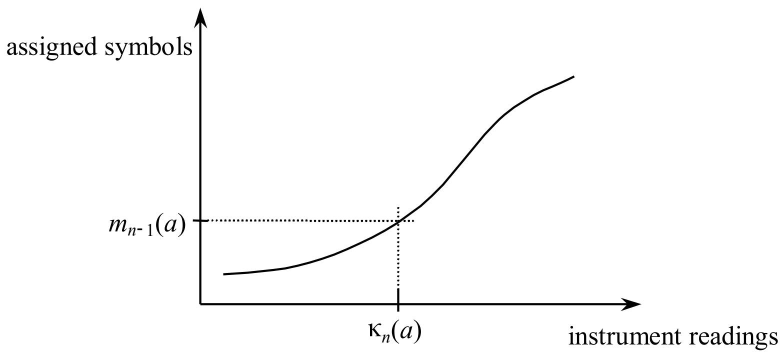
* calibration of the measuring system in reference to the standard set element a:
$\forall a \in A_{n-1}$, $m_n(a') = m_n(\kappa_n(a)) = m_{n-1}(a) = s$:



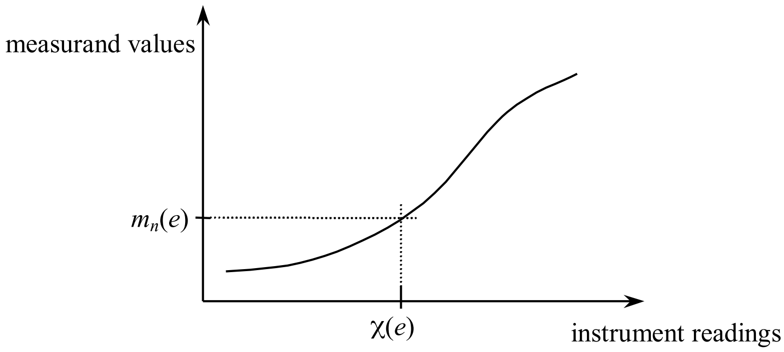
* measurement of the thing state e by means of the calibrated measuring system:
$\forall e \in E$, $s = m(e) = m_n(\chi(e))$:



Calibration can be then meaningfully expressed as the operation that empirically establishes a relation between 'instrument readings', i.e. MS states, and symbols assigned to elements of the standard set: for each element a of the standard set A_{n-1} with which the MS is put in interaction, a couple <instrument reading, assigned symbol>=<$\kappa_n(a)$,$m_{n-1}(a)$> is defined. The 'calibration function' m_n is extensionally obtained by the set of such couples, that can be visualized in a *calibration diagram* as a curve that is parametric with respect to the elements $a \in A_{n-1}$:



Once available, the same diagram can be applied to measurement, through a suitable re-interpretation of its meaning:



As a synthesis, once the MS has been calibrated:

1. the thing in its current state e is put in interaction with the MS, as formalized by the function χ; the empirical result $\chi(e)$ is a (non-calibrated) instrument reading;

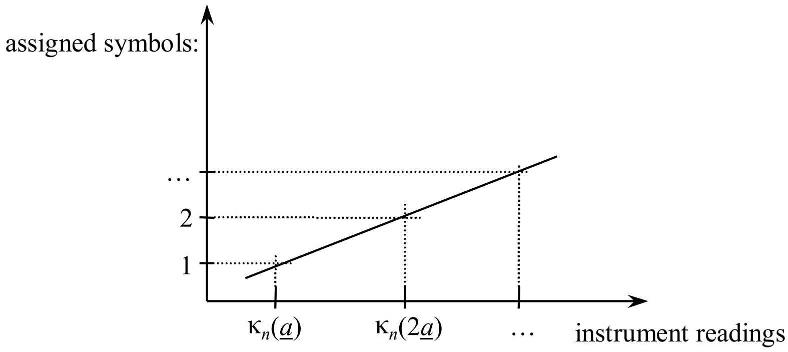
2. the instrument reading is mapped into a measurand value, as formalized by the calibration function m_n.

This procedure is based on the hypothesis of the MS *stability* in the period between the calibration and the measurement: if a given thing state e cannot be empirically distinguished from a given standard state a with respect to the measurand, then the two states must produce the same instrument reading, $\chi(e)=\kappa_n(a)$. A positive knowledge that some cause, internal or external to MS, is preventing the confirmation of this hypothesis compels to re-calibrate the MS.

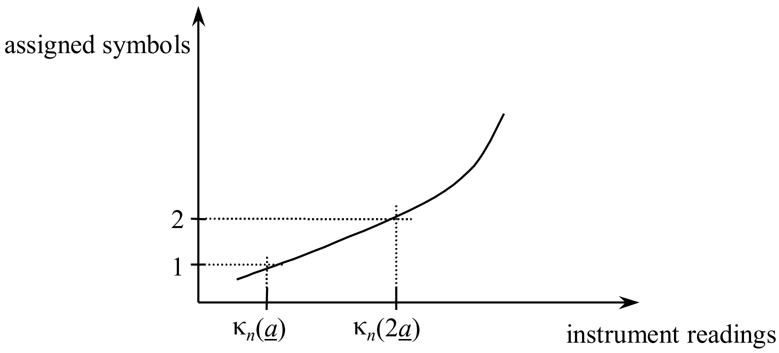
4.1. Relations between calibration and measurement

In spite of such analogies, a basic asymmetry between calibration and measurement exists.

While measurement is a single empirical operation aimed at assigning one symbol to the current thing state, calibration typically requires several assignments to be performed, in the extreme case one for each state of the MS. This is manifest in algebraically weak scales, in which to calibrate a MS the function κ_n has to be applied to each element of the standard set A_{n-1} (i.e. each element of standard set A_{n-1} must be put in empirical interaction with the MS under calibration) so that the calibration function m_n cannot be defined but in extensional way as the whole set of couples $<\kappa_n(a),m_{n-1}(a)>$. In algebraically stronger scales, in which the standard set is constructively generated by the repeated application of an operation *op* to the unit element $\underline{a}$ (so that the standard set elements can be expressed as $2\underline{a}=op(\underline{a},\underline{a})$, $3\underline{a}=op(2\underline{a},\underline{a})$, …, as previously considered), the calibration could be in principle performed as a single empirical operation by exploiting the structure of the scale, so that $m_n(\kappa_n(j\underline{a}))=jm_n(\kappa_n(\underline{a}))$ for $j=2, 3, …$ This expression can be considered the calibration-oriented version of the usual definition of ratio scale, i.e. the invariance for affine transformations $y=\alpha x$, indeed depending on the single parameter α: the only empirical operation required in this case would be the identification of the instrument reading $\kappa_n(\underline{a})$ relative to the unit element (a similar conclusion is achieved for interval scale, that would require two empirical operations for calibration, as two are the parameters on which its characteristic transformation, $y=\alpha x+\beta$, depends). Therefore:



This holds only in the ideal case, i.e. for a MS whose behavior is represented (in the case of a ratio scale) by a zero-crossing straight line in the calibration diagram. More often, a realistic identification of the MS recognizes in it several causes of non-linearity and offset from zero. The effects of such causes can be ascertained by carrying out a 'trial' calibration of the system as it would be ideal (i.e. by means of the single determination of $\kappa_n(\underline{a})$), and then revealing that for at least one standard state $j\underline{a}$ the symbol $m_n(\kappa_n(j\underline{a}))$ is different from $m_{n-1}(j\underline{a})$. This would highlight the need of an extensional definition of the calibration function, for example resulting in a diagram as:



in which the non-linear part of the curve represents the correction on the MS behavior obtained with a pointwise calibration (the example in the diagram shows a case in which, as the instrument reading values increase, the measurand values that would be obtained according to the hypothesis of linearity are lower than the ones achieved by a pointwise calibration of the MS. In those points the calibration curve correspondingly increases to compensate this effect of reduced sensitivity to high values).

It is important to remark that it is *precisely such a correction* that allows a suitably calibrated MS to preserve the scale transformations, i.e. to implement the scale-characteristic homomorphisms, in measurement.

A simple example can help us to illustrate this point. A MS for a ratio scale measurand has been calibrated as it would be ideal, i.e. by means of the single determination of the instrument reading for $\underline{a}$, so that $\kappa_n(\underline{a})$=0,5 instrument reading units (i.r.u.) and of course $m_n(\kappa_n(\underline{a}))=m_{n-1}(\underline{a})=1$ measurand units (m.u.). Afterwards the MS has been applied to three distinct thing states e_1, e_2, e_3, and the three instrument readings a'_1, a'_2, a'_3 have been correspondingly obtained, such that $a'_i=\chi(e_i)=i$ i.r.u. for i=1, 2, 3. The hypothesis of ideality would lead to the assignments $m_n(a'_i)=2i$ m.u. If, on the other hand, the MS is calibrated in a pointwise way by means of the whole standard set, it could empirically occur that $\kappa_n(2\underline{a})=1$ i.r.u. and $\kappa_n(4\underline{a})=2$ i.r.u. (as in the ideal case) but $\kappa_n(8\underline{a})=3$ i.r.u. Therefore, because of a reduced sensitivity of the MS around $8\underline{a}$, the standard set element corresponding to the instrument reading 3 is not $6\underline{a}$ but $8\underline{a}$. This implies that $e_2=2e_1$ but $e_3=4e_1$.

thing states e_i	**Instrument readings $\chi(e_i)$ (in i.r.u.)**	**measurand values $m(e_i)$ ('ideal' case) (in m.u.)**	**measurand values $m(e_i)$ ('real' case) (in m.u.)**	**thing states (ratios to e_1)**
e_1	1	2	2	1
e_2	2	4	4	2
e_3	3	6	8	4

4.2. Epistemological consequences

When adopted in measurement, a *calibrated* MS preserves the scale transformations, as shown in the last two columns of the previous table. In this perspective the mentioned purpose of the representational viewpoint to transform non-homomorphic evaluations into homomorphisms could be interpreted as an extension of the very concept of calibration to the context of behavioral sciences.

On the other hand, the epistemic characteristics of objectivity and intersubjectivity that are usually recognized to measurement results cannot be ensured by calibration alone. Complete *objectivity* means here independence of the environment, and in particular the operator, that is part of the environment, so that the measurement results convey descriptive information only on the measured thing, as it has been individuated with respect to the environment itself. Complete *intersubjectivity* means here non-ambiguity of the measurement results, so that their meaning, as expressed by the reference to the adopted RRS [5], can be integrally communicated among distinct subjects. The fulfillment of both these characteristics depends on the adoption of a proper MS, and in particular:

* objectivity implies that the MS is able to discriminate the measurand from the various influence quantities, so that the acquisition component of the MS is sensitive only to the measurand;
* intersubjectivity implies that the MS is able to refer the measurand to the primary standard, so that all the measurements expressed in terms of that standard are comparable with each other.
Therefore such characteristics cannot be considered as pre-requisites for an evaluation to be called a measurement, but a target (and plausibly *the ultimate* target) for the work of the measurement scientists and technologists [6].

A basic cause prevents the complete achievement of both objectivity and intersubjectivity of measurement results: the empirical presence of *uncertainty* in the measurement process.

5. The role of uncertainty

Calibration and measurement share a twofold nature of empirical and linguistic operations, and as such they inherit the characteristics of both. «There are certain human activities which apparently have perfect sharpness. The realm of mathematics and of logic is such a realm, par excellence. Here we have yes-no sharpness. But this yes-no sharpness is found only in the realm of things we say, as distinguished from the realm of things we do. Nothing that happens in the laboratory corresponds to the statement that a given point is either on a given line or it is not» [7].

The prominent characteristics of the *linguistic* component of calibration and measurement, as expressed by the homomorphisms $\mathbf{m}_i$, are its 'yes-no sharpness' (as far as entities such as fuzzy subsets are not used to formalize measurement results) and the conventionality of the standard set elements in the primary RRS and the associated symbols. On the other hand, the *empirical* component, as expressed by the homomorphisms $\boldsymbol{\kappa}_i$ and $\boldsymbol{\chi}$, is 'yes-no sharp' only in the ideal case, as the typical presence of some non-exactness should be realistically recognized. Uncertainties emerge in boundary cases, in which the alternative between 'yes' and 'no' cannot be sharply discriminated.

The concept of uncertainty in measurement, its role and implications have been widely covered in the scientific literature (for the Author's position, cf. [8]) and the procedures for its expression are matter of international Standards (cf. the *Guide to the expression of uncertainty in measurement* (GUM) [9]). On the other hand, the formalization of uncertainty in the context of measurement theory seems to be still an open issue. The framework that has been introduced in the previous Sections enables us to present a proposal in this regard.

In agreement with the GUM, a measurement result can be expressed as a couple <*measurand_value*, *uncertainty_value*> to make its uncertainty explicit. The GUM assumes a neutral standpoint with respect to the meaning of the term *uncertainty_value*, allowing both statistical (standard deviation of the, possibly unknown, distribution of which *measurand_value* is the estimated average: 'standard uncertainty') and set-theoretic (proportional to the width of the interval of which *measurand_value* is the center point: 'expanded uncertainty') interpretations, and both statistical ('type A') and non-statistical ('type B') methods for its evaluation. The same position will be adopted here, although it is recognized that different, and somehow more general, representations could be chosen, for example subsets, or fuzzy subsets, or probability distributions. All of these are more widely applicable than the representation suggested by the GUM, that can be employed only for algebraically strong scales, in which the concept of standard deviation is meaningful.

5.1. Uncertain calibration and measurement

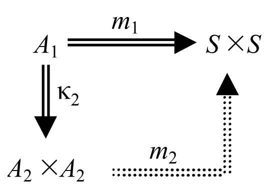
All the *empirical* functions involved in both calibration and measurement, i.e. the standard comparison functions κ_i and the thing comparison function χ, are subject to uncertainty. Their results should be accordingly expressed as couples <*measurand_value*, *uncertainty_value*>. Therefore the diagrams introduced in Section 3.2 require to be generalized as follows (double arrows indicate 'uncertain mappings'; the reader will note that an overloaded notation is adopted in the following, the same symbols referring to both non-uncertain and uncertain mappings):

1. assignment of the primary standard:

$$A_1 \xRightarrow{m_1} S \times S$$

so that $\forall a \in A_1$, $m_1(a)=\langle s,u(s)\rangle$, being s and $u(s)$ the symbols associated with the element a and its related uncertainty respectively. The term $u(s)$ highlights the presence of the so-called *intrinsic uncertainty*, i.e. an uncertainty pertaining to the very definition of the measurand and its materialization by means of the chosen primary standard. Such terms $u(s)$ (in principle one for each element of the standard set as it is mapped into S) are typically estimated on the basis of the operator's knowledge on the measurand;

2. calibration of the secondary standard:



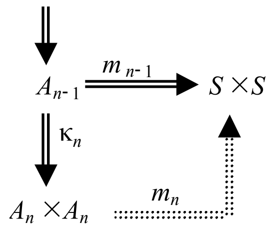
Two distinct sources of uncertainty are present in principle here: in combination with the intrinsic uncertainty of the primary standard, the effects of a non-ideal behavior of the empirical component of the calibration must be taken into account, this second component being what has been classically called the 'instrumental error' [10]. If such an error is null, i.e. in the case of ideal behavior of the system under calibration, a single instrument reading is obtained from any element of the standard set, so that $\forall a \in A_1$, $m_2(a')=m_2(\kappa_2(a))=m_1(a)=<s,u(s)>$. In this case the uncertainty of the mapping κ_2 accounts for the fact that more than one instrument reading can be empirically obtained from any single element of the standard set, so that $\forall a \in A_1$, $\kappa_2(a)=<a',u(a')>$.

In the general case, the conjoint presence of intrinsic and instrumental uncertainty must be assumed, and the issue arises of how to combine them to keep into account their contribution to the global uncertainty (for the sake of simplicity a set-theoretic interpretation will be assumed in the following discussion, so that any couple $<x,u(x)>$ identifies an interval whose center and half width are x and $ku(x)$ respectively, for a given $k>0$. A general treatment of the possible rules of combination between intrinsic and instrumental uncertainty, independent of any specific interpretation for the couples <*measurand_value*, *uncertainty_value*>, is beyond the scope of this paper, if even possible. For a discussion on this point see [8]).

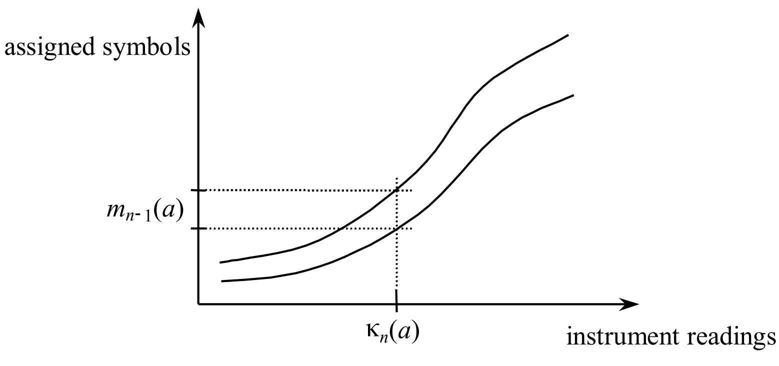
The instrumental uncertainty $u(a')$, as obtained by the function κ_2, is still negligible with respect to the intrinsic uncertainty $u(s)$ whenever $\forall a_i,a_j \in A_1$, if $m_1(a_i) \cap m_1(a_j)=\varnothing$ then $\kappa_2(a_i) \cap \kappa_2(a_j)=\varnothing$, i.e. the intervals of the instrument readings corresponding to distinguishable standard set elements do not overlap. In this case the instrument readings can be partitioned into 'groups of mutually compatible values', i.e. distinct intervals whose elements are pairwise disjoint, so that the compatibility relation among elements belonging to the same interval is transitive, and then actually an equivalence [11]. This condition is obtained by a suitable coupling of the chosen primary standard and the instrument to be calibrated by means of it, for example by reducing the number of elements in the standard set and correspondingly increasing their mutual separation, a choice highlighting the trade-off between measurand specificity and uncertainty, the two components by which the measurement information is expressed [8]. If the instrumental uncertainty is negligible with respect to the intrinsic uncertainty, the terms $m_2(a')$ can be again computed as in the ideal case: $\forall a \in A_1$, $m_2(a')=m_2(\kappa_2(a))=m_1(a)=<s,u(s)>$, i.e. the calibration leads to associate each instrument reading in the interval $\kappa_2(a)$ with the same interval $m_1(a)$.

On the other hand, in the typical experimental situation the instrumental and intrinsic uncertainties are at least comparable. As a consequence, the same instrument reading a' could be obtained from more than one standard set element. From a theoretical point of view the case is straightforward: $\forall a' \in A_2$, let $A_{1,a'}$ be the interval of A_1 of such standard set elements. Then $m_2(a')$ is computed as the set-theoretical union of the intervals $m_1(a)$ for each $a \in A_{1,a'}$. The difficulty stands in the fact that this computation requires the mapping κ_2 to be inverted: instead of applying each standard set element and observing the related instrument reading(s), the standard set elements should be found from which any given instrument reading could be generated, a strategy that is clearly cumbersome from the operative point of view;

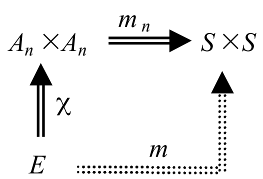
3. calibration of the MS, i.e. n-ary standard:



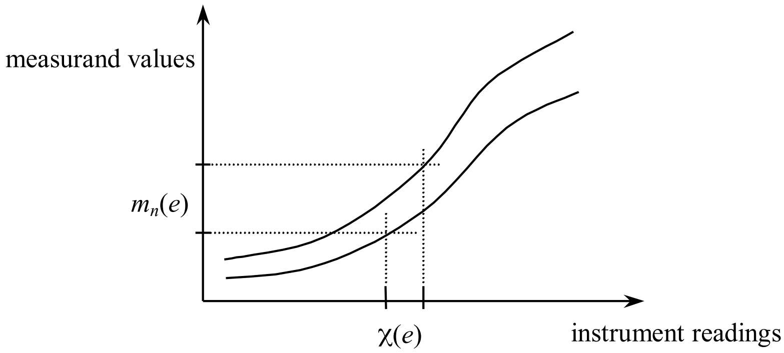
What has been said for the calibration of the secondary standard can be recursively applied to the calibration of all the standards in the traceability chain. It can be noted that each step of the chain realistically introduces some instrumental uncertainty, so that $\forall a \in A_1$ the interval $\tau(a)$ of the readings (indirectly) generated by the standard set element a in the MS is wider than the interval $\kappa_2(a)$ of the readings (directly) generated by the same standard set element in the instrument adopted as secondary standard (or, equivalently, the intervals $m_n(\tau(a))$ are more overlapped with each other than the intervals $m_1(a)$). The whole process of calibration leads to associate each MS instrument reading with an interval of symbols in S. This information can be visualized in a calibration diagram as a strip that is parametric with respect to the elements $a \in A_{n-1}$ (although from different, and non formal, premises, this is precisely the conclusion of [12]):



4. measurement:



The uncertainty identified during the calibration must be combined with the experimental uncertainty that the MS exhibits during its usage for thing comparison. Indeed any given thing state $e \in E$ can generate more than one instrument reading, the interval of such readings being formalized as a couple $\chi(e)$=<a,$u(a)$>. $\forall e \in E$, let $A_{n,e}$ be the interval of A_n of these instrument readings. Hence $m(e)$ is computed as the set-theoretical union of the intervals $m_n(a)$ for each $a \in A_{n,e}$. Again, this information is effectively visualized in a calibration diagram:



highlighting the fact that the presence of this experimental uncertainty increases the uncertainty to be assigned to the measurement results.

5.2. The implications for the representational viewpoint

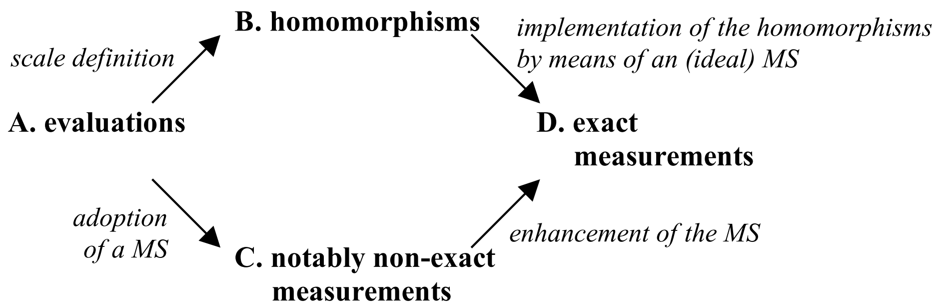
When the exactness of the measurement can be assumed, the mappings involved in MS calibration inherit the property of being relation-preserving functions, i.e. homomorphisms, in their materialization in the MS. The previous discussion shows that the relation-preservation is maintained also in the more general situation in which the instrument readings can be arranged in groups of mutually compatible values: in this case the mappings κ_i and m_i are still homomorphisms, with the trivial condition that their range set is extended to keep into account the presence of the term *uncertainty_value*. On the other hand, in presence of notable non-exactness distinguishable standard set elements generate overlapping intervals of possible instrument readings, so that the transitivity of the involved relation cannot be assured, and the mappings κ_i and m_i lose the property of being homomorphisms. This feature could be adopted as an operational criterion *to define* the concept of ‘notable non-exactness’:

a MS calibrated in reference to a primary RRS is said to introduce a notable non-exactness whenever the relations in the RRS are not preserved by the MS application.

This emphasizes the need to further detail the conceptual hierarchy that has been previously introduced:

evaluations ⊃ homomorphisms ⊃ measurements

(we are neglecting the distinction between homomorphisms and m-homomorphisms, inessential here), and that must be transformed into a 2-dimensional structure:



Once more, the diagram makes clear the distinct, and somehow opposite, paths to measurement usually followed by social and behavioral scientists (A→B and more rarely →D) vs. natural scientists and engineers (A→C and when empirically possible →D). If taken in rigorous terms, the representational viewpoint would oblige to consider that those measurements that are notably non-exact (plausibly a huge fraction of the measurements performed by natural scientists and engineers) are *not* 'measurements' at all!

6. Conclusions

In this paper the standpoint has been presented and discussed that measurement is an evaluation performed by means of a calibrated measuring system (MS). This manifestly contrasts with the representational viewpoint, that instead emphasizes scale consistency conditions.

A functional concept of MS has been formally defined, distinguishing between the ideal condition of exactness and the more general situation in which the presence of both intrinsic and instrumental uncertainty is taken into account in its consequences. It has been shown that in the former case the requirement of scale definition is empirically embedded in the MS and therefore is fulfilled by the simple usage of the MS itself, that plays the role of a relation-preserving device. On the other hand, when the instrumental uncertainty is not negligible with respect to intrinsic uncertainty the mapping materialized by the MS loses the property of relation-preservation because of the non-transitivity of the relation induced on the measurement result set by the MS application. Furthermore, in this case a complete calibration, leading to establish both a measurand value and an uncertainty degree for each instrument reading, becomes a cumbersome operation: that is why it is seldom performed, leaving to the operator the task of expressing the uncertainty on the basis of his knowledge and experience. On the other hand, the increasing diffusion of intelligent MSs, that should be able to produce complete measurement results (therefore including uncertainty estimates) in a fully automated process, stimulates the effort of defining standardized procedures to derive and express uncertainty, such as the one sketched here.

Finally, from the analysis of the interrelations between calibration and measurement the claimed objectivity and intersubjectivity of measurement have been justified in terms of the MS structure and characteristics, a valuable epistemological result to characterize measurement as a particular kind of evaluation.

A glossary of the notation used

MS:	measuring system (introduced in the Section 2.1)
RRS:	reference relational structure (Section 3.1)
$\mathbf{X}=<X,R_X>$	generic relational structure of a domain set X and a set R_X of relations on X (Section 2)
$\mathbf{m}=<m,m_R>$, $\mathbf{m}:\mathbf{X}\to\mathbf{Y}$	homomorphism from a relational structure **X** to a relational structure **Y**, such that $m:X\to Y$ and $m_R:R_X\to R_Y$ (Section 2)
$\mathbf{E}=<E,R_E>$	empirical relational structure (Section 2)
$\mathbf{S}=<S,R_S>$	symbolic relational structure (Section 2)
$\mathbf{A}_i=<A_i,R_{Ai}>$	i-ary reference relational structure (Section 3.1), so that: A_1: primary standard set $\mathbf{A}_n$: reference relational structure embedded in the measuring system
$\boldsymbol{\kappa}_i:\mathbf{A}_{i-1}\to\mathbf{A}_i$	standard comparison homomorphisms (Section 3.1)
$\boldsymbol{\tau}=\boldsymbol{\kappa}_2\to\ldots\to\boldsymbol{\kappa}_n$	traceability chain (Section 3.1)
$\mathbf{m}_i:\mathbf{A}_i\to\mathbf{S}$	calibration homomorphisms (Section 3.1)
$\chi:E\to A_n$	thing comparison function (Section 3.1) and: $\boldsymbol{\chi}:\mathbf{E}\to\mathbf{A}_n$: corresponding homomorphism (Section 3.2)
$m:E\to S$	symbol selection function (Section 3.2) and: $\mathbf{m}:\mathbf{E}\to\mathbf{S}$: corresponding homomorphism (Section 3.2)
$<s,u(s)>$	measurement result, being s the measurand value and $u(s)$ its uncertainty evaluation (Section 5.1)

References

1. Krantz, D., Luce, R., Suppes, P. and Tversky, A., *Foundations of measurement*. Academic Press, New York, Vol.1 (1971), Vol.2 (1989), Vol.3 (1990).
2. Roberts, F., *Measurement theory*. Addison−Wesley, Reading, 1979.
3. Finkelstein, L. and Leaning, M., A review of the fundamental concepts of measurement. *Measurement*, 1984, **2** (1), 25-34.
4. Mari, L., The meaning of "quantity" in measurement. *Measurement*, 1996, **17** (2), 127-138.
5. Mari, L., Notes towards a qualitative analysis of information in measurement results. *Measurement*, 1999, **25** (3), 183-192.
6. Mari, L., The role of determination and assignment in measurement. *Measurement*, 1997, **21** (3), 79-90.
7. Bridgman P.W., How much rigor is possible in physics?. In: Henkin et al. (eds.), *The Axiomatic method*, North-Holland, 1959.
8. Mari, L., Zingales, G., Uncertainty in measurement science. In: Proceedings of the IMEKO TC7 International workshop on advances in measurement science, Kyoto, 1999.
9. BIPM, ISO, et al., *Guide to the expression of uncertainty in measurement*. International Organization of Standardization, Geneva, 1993.
10. Michelini R.C., Rossi G.B., Intrinsic uncertainty in measurement: an interpretation model for metrological standards. *Measurement*, 1990, **8** (3), 103-108.
11. Mari L., On the comparison of inexact measurement results. *Measurement*, 1991, **9** (4), 157-162.
12. UNI, Misure e misurazioni. Termini e definizioni fondamentali. Italian Standard UNI 4546, 1984.