Synthese (2010) 175:123–149
DOI 10.1007/s11229-009-9466-3

Outline of a general model of measurement

Aldo Frigerio · Alessandro Giordani · Luca Mari

Received: 4 November 2008 / Accepted: 4 February 2009 / Published online: 28 February 2009

Abstract Measurement is a process aimed at acquiring and codifying information about properties of empirical entities. In this paper we provide an interpretation of such a process comparing it with what is nowadays considered the standard measurement theory, i.e., representational theory of measurement. It is maintained here that this theory has its own merits but it is incomplete and too abstract, its main weakness being the scant attention reserved to the empirical side of measurement, i.e., to measurement systems and to the ways in which the interactions of such systems with the entities under measurement provide a structure to an empirical domain. In particular it is claimed that (1) it is on the ground of the interaction with a measurement system that a partition can be induced on the domain of entities under measurement and that relations among such entities can be established, and that (2) it is the usage of measurement systems that guarantees a degree of objectivity and intersubjectivity to measurement results. As modeled in this paper, measurement systems link the abstract theory of measuring, as developed in representational terms, and the practice of measuring, as coded in standard documents such as the *International Vocabulary of Metrology*.

Keywords Measurement · Representational measurement theory · Measurement systems · Objectivity · Intersubjectivity

A. Frigerio (✉) · A. Giordani
Università Cattolica, Milan, Italy
e-mail: aldo.frigerio@unicatt.it

L. Mari
Università Carlo Cattaneo, Castellanza (Mi), Italy

1 Introduction

A general model of measurement can be aimed at: (1) providing a sound interpretation of measurement as structured process; (2) identifying the ontological conditions to be fulfilled for measurement to be possible; (3) identifying the epistemic conditions to be fulfilled for measurement results to be able to justify empirical assertions. In this paper we are mostly concerned with issue (1), assumed to be the crucial problem to be solved to pursue in facing questions (2) and (3).

Measurement can be preliminarily characterized as a fundamental process aimed at acquiring and codifying information about an entity[1] that will be called here the *system under measurement*, SuM for short. This process is commonly interpreted in functional terms as a production process, accomplished by means of a *measurement system* (MS), whose input is the SuM and whose output is a piece of information, generally expressed in numerical terms, about a SuM property. As a consequence, the task of providing a general model of measurement turns into the one of characterizing the just mentioned process.

An analysis of the concept of property is outside the scope of this paper. Accordingly, we will use "property" as a primitive term, generally denoting any empirically determinable characteristic of a SuM. Moreover, we will make use of the following definitions.[2]

- *property value*: information expressing the determination of a given property
- *measurement*: process of experimentally obtaining a property value
- *measurement procedure*: detailed description of a measurement
- *measurand*: property intended to be measured
- *measurement result*: property value being attributed to a measurand by means of a measurement
- *measurement report*: measurement result together with any other information relevant for properly interpreting the measurement result itself.

Once a measurand is selected, a process of measurement can be then structurally described as:

system under measurement → measurement system → measurement result

where the MS produces a measurement result by properly interacting with the SuM.

Any SuM can be taken into account as the instance of a class of SuMs, for all of them the same measurand is assumed to be measurable. When provided with a set of relations between its elements, this class is called a *relational system* (RS), and specifically an *empirical* RS. Hence, measurement can be functionally formalized, in its nature of process, as a mapping assigning numbers to elements of an empirical RS, in such a way that the relations between the elements in the empirical RS are preserved

[1] This term is to be intended as a general term referring to any static or dynamic system whose existence is not under discussion here.

[2] These definitions are loosely inspired to the *International Vocabulary of Metrology* (JCGM 2008a).

by relations between numbers in a numerical RS, the mapping being thus a morphism from the empirical RS to the numerical one.

empirical relational system ⟶ measurement process ⟶ numerical relational system

This is the model underlying the so-called representational standpoint of measurement theory (RTM), considered nowadays the *standard* measurement theory, in particular in social sciences. Initially proposed by Scott and Suppes (Scott and Suppes 1958), RTM has found its more extensive exposition in the monumental *Foundations of Measurement* (Krantz et al. 1971; Suppes et al. 1990; Luce et al. 1990).[3] According to representational theorists to measure *is* to construct a representation of an empirical RS to a numerical RS, under the hypothesis that relations in the empirical RS are somehow observable. By means of measurement objects in the empirical RS are associated to numbers in such a way that relations among numbers have to represent the observed relations among objects, and a numerical relation instance must correspond to each empirical relation instance.

Although its merits cannot be overestimated, we claim that RTM has some weaknesses, that can be synthesized in the following three points.

(i) To measure does not mean to associate empirical objects with numbers, but, more generally, with information entities. Only when objects can be operated by a compositional operation, the information associated with them can be expressed in numerical terms.
(ii) The association between objects and information entities is only the final stage of measurement, in itself an empirical process that requires the SuM to interact with a MS. The basic result of this interaction is the embedding of the SuM in an empirical structure: as a consequence, empirical relation instances need not to be observable before measurement, as instead RTM requires.
(iii) The epistemic features that characterize measurement with respect to generic evaluation—that we claim (see Mari 2003, 2007) to be *objectivity*, i.e., measurement results convey information on the considered SuM and not the surrounding environment, and *intersubjectivity*, i.e., measurement results can be interpreted in the same way by different subjects—cannot be reduced to the concept of representation, as instead assumed by RTM.

While the first point is grounded on the well known conception of multiple scale types, as introduced in Stevens (1946), the second point is much more critical, and accounts for the fact that RTM is rarely employed in physical metrology (as a cogent example, RTM is not even quoted in the two general documents on metrology co-authored by the main standardization bodies—the *International Vocabulary of Metrology—Basic and General Concepts and Associated Terms* (VIM: JCGM 2008a) and the *Guide to the Expression of Uncertainty in Measurement* (GUM: JCGM 2008b). The alleged reason of this failure is that measurement of physical quantities does not involve (or at least does not involve anymore) fundamental problems, which should require a theory (the measurement problem in quantum physics is a well-known exception; indeed,

[3] See also Suppes and Zinnes (1963); Pfanzagl (1968); Narens (1985, 2002) for further views on the topic.

in the current literature on physics the occurrences of "measurement" and "quantum mechanics" are strongly correlated). We claim that this reason is incomplete, if not simply false: in this paper we will instead support the position that RTM is *too abstract* to be useful in a scientific context where since a long time measurement instrumentation has been designed, built and properly used.

Our underlying goal is to characterize measurement in a way that does justice to its acknowledged epistemic role, so to explicitly justify the claim of objectivity and intersubjectivity of its results, as stated in the third point, at the same time highlighting the conditions that make such a characterization applicable both in physical metrology and in the context of the so-called *soft measurement* (see Finkelstein 2003, 2005).

This paper is structured as follows. In Sect. 2 the main aspects of RTM are illustrated, and its merits and shortcomings discussed. In the following two sections our proposal of a measurement model is outlined: in Sect. 3 the model is presented in a 'private' version (characterizing "pre-measurement", as it could be called), which accounts for objectivity conditions; in Sect. 4 such a model is extended to a 'public' version, taking into account also intersubjectivity. Sect. 5 contains some concluding remarks.

2 Account of the algebraic interpretation of measurement

Representational theories of measurement (RTM) are based on what can be called an algebraic interpretation of measurement. According to such an interpretation measurement is a process of construction of a numerical structure representing an empirical structure. To make the paper self contained, the basic features of RTM are briefly reviewed here, before analyzing what we deem are the main merits and flaws of this standpoint.

2.1 Basic definitions of RTM

First of all, let us introduce some basic definitions.

Definition 2.1 (*Relational system and relational system type*) A relational system (RS) is a pair $\boldsymbol{A} = \langle A, \langle R_i \rangle_{i \in I} \rangle$ where: (i) A is a set of elements, called the RS domain; (ii) $\langle R_i \rangle_{i \in I}$ is an ordered set of relations on A, indexed by a set I. The RS type is an ordered set of natural numbers $\langle n_i \rangle_{i \in I}$, being n_i the adicity of the relation R_i.

RTM exploits the formal construct of RS by interpreting it in terms of empirical and numerical RSs.

Definition 2.2 (*Empirical RS*) An empirical RS is a RS whose domain is a set of empirically accessible objects[4] (physical objects, events, processes, phenomena, etc.), an

[4] This definition assumes a straightforward solution to the ontological problem about the time-persistence of SuMs: each SuM is instantaneously characterized as an empirical object, i.e., empirical objects coincide with states of SuMs. For the sake of simplicity we will maintain this position here, and we will suppose the elements of the empirical RS domains to be objects, both different states of the same SuM (in particular for characterizing the SuM dynamics) and states of different SuMs (in particular for synchronously comparing SuMs).

object being said empirically accessible if it can interact with an observer or a measurement system (MS: this concept will be further characterized in the following), and if the relations it bears with other objects can be ascertained by means of some of these interactions.

Definition 2.3 (*Numerical RS*) A numerical RS is a RS whose domain is a set of mathematical objects, usually elements of Re^n, where Re is the set of real numbers and n is a natural number.

Definition 2.4 (*Scale*) A scale is a triple $\langle \boldsymbol{E}, \boldsymbol{N}, h \rangle$ where:

(i) $\boldsymbol{E}$ is an empirical RS;
(ii) $\boldsymbol{N}$ is a numerical RS;
(iii) the RS type of $\boldsymbol{E}$ and $\boldsymbol{N}$ is the same;
(iv) h is a (homo)morphism from $\boldsymbol{E}$ into $\boldsymbol{N}$.

Let us further recall that:

h is a monomorphism iff it is an injective homomorphism;
h is an isomorphism iff it is a bijective homomorphism.

A concept of *scale type* is obtained by taking into account the relative uniqueness of the morphism h, i.e., the set of permissible transformations of a scale. As a well known example, a ratio scale is unique up to similarity transformations (Stevens 1946).

Definition 2.5 (*Measurement*) Measurement consists in constructing a measurement scale, i.e., in defining a morphism on the basis of which an empirical RS is represented by a numerical RS, so that the information about the SuM is encoded by numbers.

Hence (1) a measurement process associates some objects of the domain with numbers, so that (2) relations among such objects are represented by suitable relations among numbers.

From the mathematical point of view, measurement is based on a theory proving that some relations among empirical entities can be represented by some relations among numbers. Therefore, the algebraic interpretation of measurement is aimed at ascertaining existence and uniqueness conditions on the construction of a scale, i.e., aimed at: (1) establishing whether there exists a representation of the empirical RS into a numerical RS, and (2) establishing the uniqueness of the representation, i.e., identifying which set of scales the representation gives rise to.

2.2 Paradigmatic examples

To better understand the previous definitions and the logic underlying the algebraic interpretation of measurement, two well-known examples are presented here, concerning respectively measurement of comparative quantities and extensive quantities, associated to ordinal scales and ratio scales respectively.

Definition 2.6 (*Comparative RS*) Let $\boldsymbol{E} = \langle E, \langle \blacktriangleleft \rangle \rangle$ be an empirical RS of type $\langle 2 \rangle$. $\boldsymbol{E}$ is a comparative RS iff it can be mapped into $\boldsymbol{N} = \langle Re, \langle \leq \rangle \rangle$, i.e., iff a morphism h exists, such that:

(i) $h : E \rightarrow Re$
(ii) $\forall x, y \in E(x \blacktriangleleft y \rightarrow h(x) \leq h(y))$

An ordinal scale measurement theory based on the algebraic interpretation of measurement, as a theory of scale constructability, identifies the sufficient conditions (or sometimes, as in the case of comparative RSs, the necessary and sufficient conditions) to represent the elements of E by means of elements of Re so that the instances of the relation $\blacktriangleleft$ are preserved by corresponding instances of $\leq$. Such conditions are:

C1: $\forall x, y \in E(x \blacktriangleleft y \vee y \blacktriangleleft x)$ totality
C2: $\forall x, y \in E(x \blacktriangleleft y \wedge y \blacktriangleleft z \rightarrow x \blacktriangleleft z)$ transitivity

Theorem 2.1 *a RS* $E = \langle E, \langle \blacktriangleleft \rangle \rangle$, *such that* $\blacktriangleleft$ *satisfies* **C1** *e* **C2**, *is a comparative RS.*

The proof of the theorem is not difficult (Krantz et al. 1971, ch.1).

Step 1: *the finest equivalence relation induced by* $\blacktriangleleft$ *is defined:* $x \approx y := x \blacktriangleleft y \wedge y \blacktriangleleft x$.
Step 2: *the quotient RS modulo* $\approx$ *is introduced:* $E/\approx := \langle E/\approx, \langle \blacktriangleleft^* \rangle \rangle$, *where* $[x]_\approx \blacktriangleleft^* [y]_\approx$ *iff* $x \blacktriangleleft y$. *The definition of* $\blacktriangleleft^*$ *is consistent, being* $\approx$ *a congruence in* $\boldsymbol{E}$, *i.e., an equivalence relation preserving relations between equivalent elements of the RS domain.*
Step 3: *a proof that* $\boldsymbol{E}/\approx$ *can be embedded in* $\boldsymbol{N}$ *is provided.*
The proof is standard, deriving essentially from the proof that every total dense order is isomorphic to one of the following interval in the set of rational numbers: (0,1), (0,1], [0,1), [0,1]. Dropping the condition of density, the proof continues to guarantee the injectivity of the map.

Definition 2.7 (*Extensive RS*) Let $\boldsymbol{E} = \langle E, \langle \blacktriangleleft, \bullet \rangle \rangle$ be an empirical RS of type $\langle 2, 3 \rangle$. $\boldsymbol{E}$ is an extensive RS iff it can be mapped into $\boldsymbol{N} = \langle Re, \langle \leq, + \rangle \rangle$, i.e., iff a morphism h exists, such that:

(i) $h : E \rightarrow Re$
(ii) $\forall x, y \in E\ \exists! z \in E(\bullet(x, y, z))$, so that $\bullet$ can be interpreted as a dyadic operator and written in infix form, $z \approx x \bullet y$
(iii) $\forall x, y \in E(x \blacktriangleleft y \rightarrow h(x) \leq h(y))$
(iv) $\forall x, y \in E(h(x \bullet y) = h(x) + h(y))$

A set of sufficient conditions for the mapping to be possible is the following:

E1: $\blacktriangleleft$ is a total order
E2: $\forall x, y \in E(x \bullet y \approx y \bullet x)$ commutativity
E3: $\forall x, y, z \in E((x \bullet y) \bullet z \approx x \bullet (y \bullet z))$ associativity
E4: $\forall x, y, z \in E(x \blacktriangleleft y \rightarrow (x \bullet z) \blacktriangleleft (y \bullet z))$ monotony
E5: $\forall x, y \in E(x \blacktriangleleft (x \bullet y))$ positivity
E6: $\forall x, y \in E(x \blacktriangleleft y \rightarrow \exists z \in E(x \bullet z \approx y))$ solubility
E7: $\forall x, y \in E \exists n \in N(x \blacktriangleleft ny)$ Archimedean axiom

where in formulating the Archimedean axiom, that states that the elements of the RS domain are commensurable, an inductive definition of iteration of $\bullet$ is assumed:

(i) $1x = x$
(ii) $(n+1)x = nx \bullet x$

Theorem 2.2 *a RS* $E = \langle E, \langle \blacktriangleleft, \bullet \rangle \rangle$, *where* $\blacktriangleleft$ *and* $\bullet$ *satisfy* **E1**–**E7**, *is an extensive RS.*

The first two steps of the proof are similar to the previous ones: axioms **E2** and **E4** are sufficient to guarantee that $\approx$ is a congruence in $\boldsymbol{E}$. The proof that an isomorphic embedding of $\boldsymbol{E}/\approx$ into $\boldsymbol{N}$ exists can be achieved in several ways, all resting on the basic idea that every element x in $E/\approx$ can be approximated by a series of standard sequences of 'little' elements in $E/\approx$, so that the ratio of the number of elements in a sequence approximating x and the number of elements in a sequence approximating a given unity u approaches to a given real number while the size of the elements in a sequence approaches to zero (for a detailed proof see (Birkhoff 1948, ch. 3, §1) or (Krantz et al. 1971, ch. 1)).

2.3 Credits of the algebraic interpretation of measurement

The algebraic interpretation of measurement has the important merit of identifying the sufficient conditions to represent empirical RSs by numerical RSs. The theorems proved in the context of this interpretation constrain the empirical meaningfulness (Roberts 1979) of relations applied to the numbers associated to the objects of an empirical domain to the existence of well defined, observable relations holding among such empirical objects. For example, the "more than" relation among the numbers associated to minerals on the basis of a scratch comparison, i.e., according to the so-called Mohs' hardness, can be meaningfully applied only if the scratching relation satisfies the conditions of totality and transitivity.

A further evidence of the significance of the algebraic interpretation of measurement is obtained by considering the empiricist standpoint to measurement, characterized as a decisive topic of a taxonomy distinguishing classificatory, comparative and quantitative concepts, and extensive and intensive concepts (see Hempel 1952, 1965; Carnap 1966). As an example, let us consider the introduction of comparative concepts, as defined by the previously stated axioms **C1** and **C2**. An uncritical analysis would simply ascribe to any comparative concept the same properties holding for the numerical concept. Still, under which conditions and restrictions is it possible to justify this ascription? Thanks to the algebraic interpretation of measurement the answer to this epistemological problem is straightforward, as it is derived by a representation theorem.

As a synthesis, RTM is based on the following basic ideas.

(1) SuMs are modeled as elements of empirical RSs $\langle E, \langle R_i \rangle_{i \in I} \rangle$, so that:
  (1.1) elements in E are intended to be empirically accessible;
  (1.2) relations in $\langle R_i \rangle$ are intended to be empirically decidable.
(2) Axioms about relations in $\langle R_i \rangle$ are intended to be:
  (2.1) true about the set of SuMs modeled by $\langle E, \langle R_i \rangle_{i \in I} \rangle$
  (2.2) as far as possible empirically checkable with respect to the considered SuMs.

(3) Measurement results are modeled as elements of numerical RSs $\langle N, \langle S_i \rangle_{i \in I} \rangle$:
  (3.1) whose elements are typically elements of Re^n;
  (3.2) whose relations are specified on the basis of hypotheses on the SuM.
(4) A measurement of $\langle E, \langle R_i \rangle_{i \in I} \rangle$ is given iff the following conditions are fulfilled:
  (4.1) the existence of a morphism linking $\langle E, \langle R_i \rangle_{i \in I} \rangle$ to $\langle N, \langle S_i \rangle_{i \in I} \rangle$ is proved;
  (4.2) the uniqueness of the morphism up to a class of transformations is proved.

Hence, the modeling process leading to $\langle E, \langle R_i \rangle_{i \in I} \rangle$ is viewed as an activity of idealization performed on a set of SuMs, aimed at abstracting away what is considered irrelevant and identifying the relations in $\langle R_i \rangle$ and the axioms determining how such relations are to be thought of. Such axioms usually allow the construction of standard sequences, i.e., sequences of equi-spaced elements used to determine the morphism linking the empirical RS to the numerical RS.

As a process of scale construction, measurement is then representable as:

$$\begin{tikzcd} \boldsymbol{E} \arrow[rr] \arrow[dr] & & \boldsymbol{N} \\ & \boldsymbol{E}/\!\approx \arrow[ur] & \end{tikzcd}$$

Step 1: the empirical RS $\boldsymbol{E}$ is mapped to the quotient RS $\boldsymbol{E}/\!\approx$ obtained by the finest equivalence relation defined on the basis of the sequence $\langle R_i \rangle$. In this step a homomorphism linking $\boldsymbol{E}$ to $\boldsymbol{E}/\!\approx$ is then defined.
Step 2: the quotient RS constructed in Step 1 is mapped to a numerical RS $\boldsymbol{N}$. In this step, the crucial one, $\boldsymbol{E}/\!\approx$ is then embedded in $\boldsymbol{N}$.
Step 3: the empirical RS is mapped to the numerical RS. In this step, the homomorphism from $\boldsymbol{E}$ to $\boldsymbol{E}/\!\approx$ and the embedding of $\boldsymbol{E}/\!\approx$ in $\boldsymbol{N}$ are composed to define a homomorphism from $\boldsymbol{E}$ to $\boldsymbol{N}$.
Step 4: the transformation group of the set of embeddings of $\boldsymbol{E}/\!\approx$ into $\boldsymbol{N}$ is determined, thus giving the conditions under which such embedding are to be considered unique.

2.4 Shortcomings of the algebraic interpretation of measurement

We claim that the algebraic interpretation of measurement is just a part, although an essential one, of a more general model of measurement. Some justifications are provided here to the thesis that the algebraic interpretation cannot be the whole story on measurement.
First drawback: the approach is too abstract.

The first critical issue for the algebraic interpretation is its possibility to be completely disjointed from the practice of measurement: it is not required to design and perform any concrete measurement process to be able to progress in the algebraic theory. In the extreme case, even if human beings had not ever concretely measured, the possibility of representation theorems remains unaltered, also because the condition for the elements of the domain E to be empirical objects is extrinsic to the theory. As a consequence, the basic problems in the practice of measurement—such as those concerning the design, construction and proper operation (including calibration) of measuring systems—are simply disregarded here. As already remarked, this is a plausible reason why such an approach is not adopted in the field of physical metrology, where instead such problems are pivotal.

Second drawback: the conditions of measurability seem to be too strong to be necessary.

An empirical RS is specified by selecting a set of objects, a set of relations, and a set of axioms on the relations. If the selected objects are to be interpreted as SuMs, the conditions stated by the axioms of an empirical RS are often too strong to be necessary conditions of measurement. For example, in defining extensive RSs, a version of the axiom of solubility is to be introduced, but hardly any set of empirical objects can be a model for such an axiom, since it implies the existence of a sequence of arbitrarily small quantities.[5]

Third drawback: the conditions of measurability seem to be too weak to be sufficient.

An evidence for the abstract character of the algebraic interpretation is provided by the definition of measurement given in terms of representation. According to this definition measuring is basically representing, so that measurability only requires the mere existence of a numerical RS by which an empirical RS can be represented. This is a very weak condition. Indeed, it follows from the definition that every empirical RS is measurable, in virtue of the fact that from a set-theoretical point of view the following three conditions are equivalent:

(i) $\boldsymbol{E} = \langle E, \langle R_i \rangle_{i \in I} \rangle$ is representable by a numerical RS $\boldsymbol{N} = \langle Re^n, \langle S_i \rangle_{i \in I} \rangle$;
(ii) $\boldsymbol{E} = \langle E, \langle R_i \rangle_{i \in I} \rangle$ is representable by a numerical RS $\boldsymbol{N} = \langle Re, \langle S_i \rangle_{i \in I} \rangle$;
(iii) $|E| \leq |Re|$,

where $|X|$ denotes the cardinality of X. This consequence, originally noticed by Scott and Suppes (1958, pp. 116–117), seems to be too problematic for assuming representability as a sufficient condition of measurability, at least if one supposes that measurability is, or at least involves, *empirical* possibility of measurement. Hence, measurability cannot be characterized only in terms of representability, without specifying any other condition on the involved systems.[6]

Fourth drawback, no distinction is made on the numerical RS.

Elements of an empirical RS are mapped by measurement into elements of a numerical RS, whose domain is, in the simplest case, a subset of the set of real numbers. The function played by numbers is different in different scale types, e.g., in nominal and ordinal scales with respect to ratio scales. For, in nominal and ordinal scales numbers are immaterial and can be substituted by any kind of object or sign. On the contrary, in ratio scales numbers are essential, since they convey specific information about the ratio (a numerical concept, indeed) of quantities characterizing the SuMs. Thus, even if the algebraic interpretation is able to unify different measurement scale types under a broad concept, this way of conceiving measurement leads to underestimate important problems in both the theory and the practice of measurement.

[5] A significant exception is constituted by the case of finite structures used to construct finite standard sequences, as defined, e.g., by Suppes (2006). The theory of finite standard sequences constitutes plausibly the best starting point to incorporate RTM in the general model proposed here.

[6] A similar point is highlighted in Niederée (1992) and in Narens (2002, ch.153).

In the next two sections we try to overcome these problems and to supply a richer definition of measurement by introducing a theoretical concept of measurement system, defined as a system able to interact with SuMs and provided with instructions specifying how such interaction must be performed and interpreted. In interacting with SuMs, a MS provides the empirical domain with a set of relations, thus generating an empirical RS, while in interpreting the interactions such an empirical RS is connected to a symbolic RS (not necessarily a numerical RS) in a process we will call *pre-measurement*. By properly designing and operating with a MS, the objectivity of the results is (to some degree) obtained, but still their interpretation cannot be guaranteed to be unique among different subjects. That is why pre-measurement is in fact a *private* process. Whenever the further requirement of MS calibration by a traceable reference is enforced, the process becomes *public*, because of the intersubjectivity of its results, and it can be considered a *measurement*. Hence, our thesis is that measurement is characterized as a (homomorphic) mapping of empirical objects into symbols performed by means of a MS, so to guarantee (to some degree) the objectivity and the intersubjectivity of such results.

3 Pre-measurement

We propose to characterize *pre-measurement* as an experimental process aimed at defining a set of relations in a domain E of empirical objects by means of a *pre-measurement system* (PMS),[7] operatively defined as a 4-tuple $\langle M, \mathrm{P1}, \mathrm{P2}, \mathrm{P3}\rangle$, where:

(i) M is a system able to interact with an intended set of SuMs in given states;
(ii) P1–P3 are the three parts of a user's manual in which the operating instructions for using M, i.e., its related measurement procedure, are specified.[8]

A pre-measurement is performed by properly resetting M before its usage, then having M interact with the given SuM. This interaction is supposed to lead to a specific state of the coupled system, here denoted M■SuM, typically but not necessarily by a state transition of M, in this case thought of as a dynamical system. P1–P3 are the specifications aimed at properly setting up and performing such an interaction and interpreting its result. P1–P3 are defined as follows:

[7] We are introducing here the term "pre-measurement (system)" to emphasize a subset of features of a customary measurement (system). Roughly speaking, a PMS can be thought of as a non-calibrated measurement system. The concepts of system and system state will be used here as follows. A system is an object, typically but not necessarily a physical one, that can be described using a non-empty set S of states, where a state, from an empirical point of view, is simply a set of discernible properties of the object and, from an abstract point of view, is a point in a given abstract space, modeling the set S itself. If a system is described by a set including more than one state, an initial state can be selected, in such a way that a concept of system evolution can be introduced, described as a transition from the initial state to another state in S. While such a characterization of a system can be straightforwardly developed into the mathematical theory of dynamical systems, the adoption of this theory is not required for understanding what a PMS is nor for describing pre-measurements performed by means of PMSs.

[8] A PMS can be chosen among others for several reasons. Often the choice is grounded on a model of the measurand that precedes the measurement. But sometimes the contrary is true: a quantity is discovered by means of a PMS (see Kuhn 1961). We will not further discuss this subject here.

–P1 (instructions concerning *the initial step*) specifies how M must be setup to interact with candidate SuMs, in their turn possibly prepared for such an interaction.[9] In the case M is a dynamical system, P1 identifies a subset of initial states of M, and lists the operations required to bring M to an initial state by a suitable state transition.
Hence, this Part emphasizes a *controllability*[10] requirement on M.

–P2 (instructions concerning *the transition step*) define an empirical RS $\langle U, \langle R_i \rangle_{i \in I} \rangle$ and specifies how M, as set up according to P1, must interact with any candidate SuM to generate a coupled system M■SuM, whose state is an element in the support U of $\langle U, \langle R_i \rangle_{i \in I} \rangle$.$U$ is a set of *selectable elements*, i.e., a set of possible states of M■SuM that are (i) empirically discernable and (ii) predictably determinant from the point of view of pre-measurement. P2 also lists the operations required to check whether the coupling operation results in a selectable element. In the case M is a dynamical system, P2 typically identifies as fixed elements some states of M itself, sometimes called "indications" or "readings" in the context of measurement of physical quantities, that are supposed to be obtainable by the state transition induced by the interaction of M with a SuM.
Hence, this Part emphasizes an *interactivity* and *observability* requirement on M.

–P3 (instructions concerning *the final step*) includes instructions aimed at codifying the information obtained in the transition step. P3 specifies a function mapping each selectable element of U, as defined in P2, into an element of the support of a given symbolic RS $\langle V, \langle S_i \rangle_{i \in I} \rangle$ and each empirical relation R_i of the RS $\langle U, \langle R_i \rangle_{i \in I} \rangle$ into a corresponding relational symbol S_i.
Hence, this Part emphasizes an *interpretability* requirement on M.

The *interaction* performed according to P2 can be formalized as $\kappa_{\mathrm{PMS}}: E \to U$, $u = \kappa_{\mathrm{PMS}}(e)$ (the subscript will be omitted when this does not generate ambiguities), to be read: the PMS interacted with the SuM in its (generally unknown) state $e \in E$, and $u \in U$ is the selectable element obtained from the interaction (note that the initial state of M can be omitted in this notation because it is specified by P1).

The *symbolization* performed according to P3 can be then formalized as λ_{PMS}: $U \to_V$, $V = \lambda_{\mathrm{PMS}}(u)$ (also this subscript will be omitted when appropriate), so that $v = \lambda_{\mathrm{PMS}}(\kappa_{\mathrm{PMS}}(e))$ is the pre-measurement result obtained by the application of the PMS $\langle M$,P1,P2,P3$\rangle$ to the given SuM in the state e.

3.1 Remarks on the definition.

Both the design of M and the definition of P1-P3 are generally based on a theoretical framework, i.e., a set of background theories, specifying in particular:

(i) how M and any candidate SuM can be made interacting with each other;

[9] Such an operation, aimed at improving measurement systems performance, is typically executed in electric and electronic measurements, where it is called "signal conditioning".

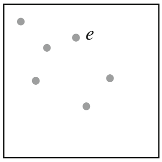
[10] The term "control" is adopted here as used in Control Theory, i.e., as a process by which the inputs to the system under consideration are manipulated to obtain the desired effect on the output of the system itself.

(ii) how to ascertain, so far as possible, whether the coupled system M■SuM is closed to other interactions;
(iii) how to take into account the effects of other interactions.

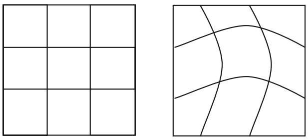
The explicit reference to a given theoretical framework is not however essential for performing pre-measurements by means of a MS, that can survive the substitution or improvement of its background theories. Still, the development of a discipline can result in a deeper understanding of the pre-measurement process, with respect to both the interaction between PMSs and SuMs and the conditions under which the coupled system can be considered closed to external influences. As a consequence, while the characterization of a PMS only requires a suitable description of M and a presentation of P1–P3, as shown in the following examples, the detailed account of a concrete pre-measurement process performed by using a given PMS is possible only within the theoretical framework of the interested discipline.

Example 1 position pre-measurement by a net.

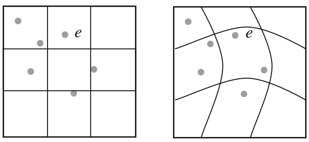
The considered empirical objects are points on a plane E, and some information on the position of a point e in E is required.



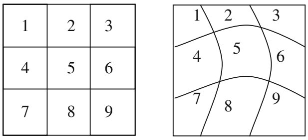
To this goal, a procedure is defined according to which a physical net M is drawn on the plane, so to have each point belonging to one and only one of the regions generated by M on E. P1 specifies how to prepare the net for its application to the plane. Two possible alternatives (PMS1 and PMS2) are as follows:



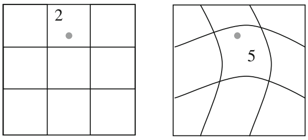
P2 specifies how the net must be drawn on the plane, identifies the meshes of the net as the selectable elements of U (in this case $\langle R_i \rangle$ is empty) and specifies how to assess in which mesh each point is contained, so that the mapping κ is interpreted as points $\rightarrow$ meshes.



Finally, P3 specifies a function λ from meshes to symbols, for example chosen in a set $V = \{1, 2, \ldots, 9\}$, as follows:



Pre-measurement of the point e position by PMS1 and PMS2 thus gives:



so that the pre-measurement result for e is $v = 2$ if PMS1 is adopted and $v = 5$ if PMS2 is adopted instead.

Example 2 hardness pre-measurement by scratching. The considered empirical objects are minerals in a set E, and some information on the hardness of an object e in E is required. To this goal, a set M of n reference minerals is used.

- P1 specifies that for every couple of reference objects of M a scratch relation R is experimentally determined, this relation being a total order (so that for any couple of reference objects either the first scratches the second or vice versa), and correspondingly arranges such reference objects in increasing order.
- P2 identifies the reference objects of M as the selectable elements of U and the total order R on M as the only relevant relation in this RS. Moreover it specifies how a scratch sequence is set up, for example by rubbing every reference object of M on the object e in the given orderly way until a scratch appears on e, so that the mapping κ is interpreted as candidate objects $\rightarrow$ reference objects.
- Finally, P3 specifies a function λ from reference objects to symbols, for example chosen in a set $V = \{1, 2, \ldots, n\}$ (any totally ordered set of n elements could be chosen in this case) so that the symbol i is assigned to the i-th reference object in the order specified in P1.

Hence, the pre-measurement result for e is i if the i-th reference object scratched e.

3.2 Differences with respect to the algebraic interpretation of measurement

Given the assumption that empirical objects $e \in E$ are states of SuMs (see footnote 3), it generally happens that $\exists e_i, e_j \in E(e_i \neq e_j \wedge \kappa(e_i) = \kappa(e_j))$, i.e., distinct objects are mapped into the same state of M■SuM. As a consequence, the application of the mapping κ_{PMS} can be meaningfully interpreted as inducing a canonical partition on the domain E, such that $e_i \approx_{\mathrm{PMS}} e_j$ iff $\kappa_{\mathrm{PMS}}(e_i) = \kappa_{\mathrm{PMS}}(e_j)$, where $\approx_{\mathrm{PMS}}$ is an equivalence relation on E formalizing a concept of mutual substitutability among the object in E with respect to the selectable elements for the given PMS. More generally, any empirically meaningful relation on the set U of selectable elements is also induced on the domain E, so that, for example in the case of ordinal relations, the results related to Definition 2.6 and Theorem 2.1 hold.[11] On the other hand, some radical differences of this situation with respect to the scenarios usually dealt with by the algebraic interpretation of measurement should not be underestimated:

(1) the assumption of observability for the relation instances in the empirical RS $\boldsymbol{E}$ *does not play any role in pre-measurement*: what it is assumed here is only the selectability of the elements of U;
(2) the assumption of existence of relations on E *does not play any role in pre-measurement*: what it is assumed here is only the existence of relations on the set U of the selectable elements for M;
(3) the assumption that the relations on E satisfy specific axioms *does not play any role in pre-measurement*: the usage of a PMS, as specified in its user's manual, is a sufficient condition to achieve measurement information about the given SuM;
(4) the assumption of existence of a morphism mapping the empirical RS $\boldsymbol{E}$ into a numerical RS *does not play any role in pre-measurement*: the constructability of such a morphism does not seem to be a necessary condition neither for defining nor for performing pre-measurement.

3.3 Objectivity of pre-measurement

The adoption of a PMS gives the fundamental benefit of guaranteeing some degree of *objectivity* to evaluation, whose results depend on the SuM and the PMS but possibly not anything else, i.e., they are in principle independent of the surrounding environment and the evaluating subject. On the other hand, the objectivity of pre-measurement critically depends on the quality of the interaction between the SuM and M, as performed in environmental conditions $h(t)$ at the time t, such conditions being usually modeled in terms of *influence quantities* (so that generally h is a vector). Indeed, it can happen that the repeated application of the same PMS to the (hypothesized) same SuM state e at different times $t_i \neq t_j$, possibly in different environmental conditions $h(t_i) \neq h(t_j)$, leads to different selectable elements, $\kappa_{t\,i}, h(t_i)(e) \neq \kappa_{t\,j}, h(t_j)(e)$. Two basic causes can be identified for this behavior:

– M is not perfectly stable in its interaction with the SuM, so that it could produce results that depend not only on the SuM but also, e.g., on some noise internal to M itself[12]. Hence, stability can be formalized by the finite-difference ratio $\Delta\kappa_t, h(t)(e)/\Delta t$ (or the partial derivative $\partial\kappa_t, h(t)(e)/\partial t$ for continuous time models), such that the less is this ratio the more M is stable. Of course, stability is usually a matter of degrees instead of a yes-no feature of the PMS: *the more a PMS is stable, the more its results are objective*;

[11] One could legitimately ask whether the structure induced by the PMS exists in the empirical domain independently of the PMS, so that the PMS itself just elicits it, or it does not preexist and is actually yielded by the PMS application. We remain neutral on this ontological issue here. The model we are presenting is also neutral on the existence of relations among the results of the coupling of the same SuM with different PMS: both Bridgman's operationalism (Bridgman 1927) and the theories claiming that the same quantity can be measured by different PMS are compatible with the standpoint presented in this paper.

[12] In physical metrology this concept is usually distinguished in long-term (e.g., days to months) stability and short-term stability, the latter being called "repeatability" of the MS.

– the MS is not perfectly selective with respect to the influence of the environment surrounding the SuM, so that it could produce results that depend not only on the SuM but also on its environment. Hence, selectivity[13] can be formalized by the finite-difference ratio $\Delta\kappa_t, h(t)(e)/\Delta h(t)$ (or the partial derivative $\partial\kappa_t, h(t)(e)/\partial h(t)$ for continuous time models), such that the less is this ratio the more M is selective. Once again, selectivity is usually a matter of degrees instead of a yes–no feature of the PMS: *the more a PMS is selective, the more its results are objective*.

Furthermore, the formalization $v = \lambda(\kappa(e))$ highlights one of the main, both conceptual and operative, problems of measurement. Pre-measurement results are obtained by symbolization of elements of U related to the coupled system M■SuM, i.e., $v = \lambda(u)$. On the other hand, such results are interpreted as expressing information on SuM states, as highlighted by $u = \kappa(e)$. The issue is that the empirical interaction formalized by the mapping κ could produce the critical side-effect of modifying the SuM state e (a situation that particularly in soft measurement is the rule more than the exception), so that the element u, and then the pre-measurement result v, would refer to such a modified state instead of the initial, unaffected, SuM state e. Two general and complementary techniques are adopted to cope with this problem: (1) PMSs are designed so to minimize their influence on SuMs, and (2) such an influence is modeled and then formalized as a correction factor that is introduced in the symbolization function λ, sometimes called "measurand reconstruction" for this reason. Let us call "non-invasiveness" this ability of a PMS to interact with SuMs while leaving their state unmodified. Hence, *the more a PMS is non-invasive, the more its results are objective*.

Pre-measurement can be thus interpreted in its turn as a fuzzy concept, in the sense that the more an evaluation process is objective, being based on a stable, selective, and non-invasive PMS, the more it is a pre-measurement. Moreover, the fact that in different fields different standards of stability, selectivity, and non-invasiveness are customary explains why for the same term "measurement" partially different acceptation are adopted, by differently putting the threshold of the acceptance criterion in the continuum from purely subjective to completely objective processes.

It should be finally noticed that, under the hypothesis of that the interaction between M and the SuM is repeatable, so that a statistical distribution of selectable elements can be obtained, objectivity of pre-measurement results can be increased by the application of several data processing, both time-domain (e.g., means) and frequency-domain (e.g., filters) based, techniques to such distributions. On the other hand, the analysis of these techniques is outside the stated aims and scope of this paper and therefore it will not be further discussed here.

Example 3 ordinal weight pre-measurement by a two pan balance. Some information on the weight of an object e in a set E is required. To this goal, M is a set of n reference objects, together with a two pan balance operated as a comparator.

[13] The reader should be aware that selectivity, i.e., the property related to the immunity to influence quantities, is not equivalent to selectability, i.e., the property according to which selectable elements of U can be identified.

- P1 specifies that, after the balance pans have been aligned and any oscillation has been stopped, for every couple of reference objects of M a weight comparison relation is experimentally determined by the balance itself, this relation being a total order (so that for any couple of reference objects, each of them put on a pan of the balance, either the left pan or the right one lowers), and correspondingly arranges such objects in increasing order.
- P2 identifies the reference objects of M as the selectable elements of U and the total order relation R on M as the unique relevant relation in this RS. Furthermore P2 specifies how a weighing process must be performed, for example by putting e on the left pan of the balance and the reference objects of M one by one in the given orderly way on the other pan until the right pan reaches a lower position than the left one, so that the mapping κ is interpreted as candidate objects $\rightarrow$ reference objects.
- Finally, P3 specifies a function λ from reference objects to symbols, for example chosen in a set $V = \{1, 2, \ldots, n\}$ (any totally ordered set of n elements could be chosen in this case) so that the symbol i is assigned to the i-th object in the sequence specified in P1.

Hence, the pre-measurement result for e is i if the i-th reference object was the first to position the right pan lower than the left one.

Example 4 ratio weight pre-measurement by a two pan balance. As in the previous example, some information on the weight of an object e in a set E is required. Once more, M is a set of n reference objects, together with a two pan balance.

- P1 specifies that, after the balance pans have been aligned, for every couple of reference objects of M a weight comparison relation is experimentally determined by the balance itself, this relation being an equivalence with a single equivalence class, i.e., all reference objects belong to the same equivalence class, so that for any couple of reference objects, each of them put on a pan of the balance, neither of the pans lowers.
- P2 identifies the sets of reference objects of M as the selectable elements of U and the relation $<$ on the numbers of elements of these sets as the only relevant relation in this RS. Moreover P2 specifies how a weighing process must be performed, for example by putting e on the left pan of the balance and one, then two, then … reference objects of M on the other pan until the right pan reaches the same or a lower position than the left one, so that the mapping κ is interpreted as candidate objects $\rightarrow$ sets of reference objects.
- Finally, P3 specifies a function λ from sets of reference objects to numerals in the set $V = \{1, 2, \ldots, n\}$ so that the symbol i is assigned to a set of i objects.

Hence, the pre-measurement result for e is i if the set of i reference objects was the first to position the right pan at the same height or lower than the left one.

Example 5 weight pre-measurement by a spring dynamometer.

As in the previous example, some information on the weight of an object e in a set E is required. In this case a dynamometer M is used, whose spring elongation is related to the applied weight force intensity.

- P1 specifies how to properly set up the dynamometer, typically by letting its spring move to the rest position.

– P2 identifies the observable elongations of the spring as the selectable elements of U and the relation $<$ on these elongations as the relevant relation in this RS. Moreover, P2 specifies how e must be suspended to the spring for obtaining a stable elongation, so that the mapping κ is interpreted as candidate objects $\rightarrow$ spring elongations.
– Finally, P3 specifies a function λ from spring elongations to numerals in the set $V = \{1, 2, \ldots, n\}$, where n is the number of individually observable elongations, so that the symbol i is assigned to the i-th elongation.

Hence, the pre-measurement result for e is i if the i-th elongation was obtained by suspending e to the dynamometer spring.

3.3.1 Remarks on Examples 1–5

In the Examples 1–5 PMSs are based on physical devices, specifically designed, built and operated with the aim of guaranteeing the objectivity of their results. From this point of view, let us discuss with some more details the PMS presented in Example 2.

Reference minerals in the set M are chosen so to be:

– *as stable as possible*, i.e., if the possible time dependence of the scratch relation R is made explicit, so that $R_t(e,u_i)$ denotes that the object e scratched the reference mineral u_i at the time t, then ideally $\forall t\ (R_t(e,u_i))$ as far as the state e is assumed to be time constant;
– *as selective as possible*, i.e., if the possible dependence of the scratch relation R on external influence quantities (e.g., environmental temperature) is made explicit, so that $R_{\boldsymbol{q}}(e,u_i)$ denotes that the object e scratched the reference mineral u_i when the measured value for the relevant influence quantities is $\boldsymbol{q}$, then ideally $\forall q\,(R_{\boldsymbol{q}}(e, u_i))$ as far as the state e is assumed to be time constant;
– *as non-invasive as possible*: in its simplest acceptation, invasiveness implies that any interaction with a reference mineral in M modifies the state of the SuM; then, for an ideally non-invasive PMS, $R_{t_1}(e, u_i) = R_{t_2}(e, u_i)$, being t_1 and t_2 the time immediately before and immediately after the scratch relation is performed, under the hypothesis of the time constancy of all relevant influence quantities in the interval $[\,t_1, t_2]$.

This characterization highlights how critical is objectivity to validate pre-measurement results. Let us assume that the (hypothetically) same PMS interacts with the (hypothetically) same SuM at the times t_1 and t_2, and two different pre-measurement results are obtained, as derived from $R_{t_1}(e, u_i) \neq R_{t_2}(e, u_i)$. How should this situation be accounted for? (i) Is it caused by the low stability of the PMS, i.e., e interacted with reference minerals that changed their state from t_1 to t_2? Or (ii) is it caused by the low selectivity of the PMS, i.e., the results of the scratch relation depends on the interaction of e with the reference minerals and the surrounding environment, that changed its state from t_1 to t_2? Or (iii) is it caused by the low non-invasiveness of the PMS, i.e., actually the SuM changed its state from t_1 to t_2, but this was only caused by its interaction with the PMS? Or (iv) is it the positive information of the SuM 'spontaneous' state change from t_1 to t_2? Or finally (v) is it the superposition of two or more of the previous causes?

The more the model of the PMS credibly supports its objectivity, the more the right answer can be assumed to be the fourth one. Vice versa, evaluation systems whose objectivity cannot be assessed lead, in general, to ambiguous results with respect to the alternatives (i–iv), and therefore to the poorly informative option (v).

A second general remark relates to the critical importance of the measurement procedure specification, i.e., elements P1-P3 in the presented model of PMS, to experimentally characterize a given measurement process. Indeed, Examples 3 and 4 highlight that the same physical device could be exploited to implement different measurement procedures. In the context of measurement of physical quantities such procedures are classified in terms of *measurement methods*.[14]

The comprehensiveness of the PMS model allows to easily particularizing it so to include such methods as categories of elements P1–P3.

Example 6 suitability for admission to university pre-measurement by multiple-choice test.

The considered empirical objects are students in a set E, and some information on their suitability for admission to a university is required. To this goal, a questionnaire M composed of n closed questions, each of them with a single right answer, is used.

- P1 specifies that the questionnaire must be submitted in its blank state, i.e., with all questions still unchecked.
- P2 identifies all possible sets of right answers as the selectable elements of U and the relation $<$ on the cardinality of such sets as the relevant relation in this RS. Moreover, P2 specifies how the questionnaire must be submitted (e.g., instructions concerning the duration of the test and the sources of information that can be possibly used), so that the mapping κ is interpreted as candidate students $\rightarrow$ sets of right answers.
- Finally, P3 specifies a function λ from sets of right answers to symbols. In the simplest case in which only the number of right answers is taken into account, symbols are chosen in the set $V = \{1, 2, \ldots, n\}$, so that the symbol i is assigned to the questionnaires with i right answers.

[14] See, for instance, the following definitions taken from *Electropedia* (IEC 2008), an excellent online reference:
–comparison (method of) measurement: "method of measurement based on the comparison of a measurand with a known quantity of the same kind";
–substitution (method of) measurement: "comparison method of measurement in which a measurand is replaced by a known quantity of the same kind, chosen in such a manner that the effects of these two values on the measuring instrument are the same";
–complementary (method of) measurement: "comparison method of measurement in which the measurand is combined with a known quantity chosen in such a manner that the sum of their values is equal to a predetermined comparison value";
–differential (method of) measurement: "comparison method of measurement, based on comparing the measurand with a quantity of the same kind having a known value only slightly different from that of the measurand, and measuring the algebraic difference between the values of these two quantities";
–null (method of) measurement: "differential method of measurement where the difference between the value of the measurand and the known value of a quantity of the same kind, with which it is compared, is brought to zero".

Hence, in this case the pre-measurement result for *e* is *i* if the *i* is the number of right answers *e* has given to the questionnaire.

3.3.2 Remarks on Example 6

The use of questionnaires as PMSs raises some problems concerning the definition of the PMS itself, the SuM, and the measurand. The case proposed in the previous example ("first case" in the lists below) is straightforward: the PMS substantially coincides with the questionnaire and the SuM with a student, while the measurand is evaluated by simply counting the right answers. On the other hand, a questionnaire could be used also, e.g., as a device supporting the interview to subjects who are asked to evaluate the quality of a product ("second case" in the lists below). By comparing these examples some ambiguities emerge, which can be removed by properly characterizing questionnaires as PMSs, or components of them.

(1) The SuM is:
   - the person who is asked to answer the questions (first case);
   - something else (second case: a product).
(2) The PMS is:
   - the questionnaire itself (first case);
   - the questionnaire together with the subject answering the questions (second case).
(3) The measurand can be:
   - measured directly (first case);
   - measured indirectly (second case, if, as it is typical, evaluated as a function of the answers given by multiple subjects).

   Moreover, there are different kinds of questionnaires that can differ with each other for the following structural aspects.

A. Relations among questions:
   - there is only one question (or there are several mutually independent ones, each of them related to a different measurand);
   - questions are interpreted as acquisition devices for the same measurand: they are chosen for redundancy reasons and for enhancing sampling frequency, and the obtained information must be synthesized statistically;
   - questions are interpreted as acquisition devices for different but not independent measurands: the obtained information must be synthesized by some data fusion logic.
B. Relations between the answering subject and the SuM:
   - the SuM is the answering subject (e.g., her competence on a given matter is under measurement);
   - the SuM and the answering subjects are distinct entities (e.g., the teaching skill of a teacher is under measurement and students are asked to perform such an evaluation);
   - the SuM includes the subject but does not coincide with it (e.g., the popularity rating of an academic course is measured and students are asked to evaluate how much the course was appreciated).

C. Relations between the questionnaire and the measurand:
   – the questionnaire measures the measurand directly (as in the first case);
   – the questionnaire measures the measurand indirectly through a functional relation (e.g., if students are asked to evaluate the teaching skill of a teacher);
   – the questionnaire measures only partly the measurand.

4 Measurement

Measurement is a process aimed at producing information interpretable in the same way by different subjects in different contexts and times, i.e., aimed at being intersubjective and therefore public in its results. In contrast, the relations among the objects determined by PMSs are meaningful independently of the mapping between empirical objects and symbols, which only formalizes the information acquired by the use of the PMS. As a consequence, such information can be interpreted only in explicit reference to the PMS that was used to obtain it: that is why pre-measurement is indeed a private process.

An apparently straightforward strategy to reach the intersubjectivity that is required to measurement is to complement the PMS user's manual with a further set of specifications, stating how *M* can be replicated in multiple instances, all of them aimed at producing the same symbolic results from the same SuM in the same state. This makes the system stability, previously introduced as a requirement for objectivity, a critical condition also for intersubjectivity. While in principle feasible, this strategy is seldom adopted, because it implies recurring peer comparisons among SuMs to assess the compatibility of their results, possibly followed by some adjustment operations. Furthermore, compatibility is not necessarily transitive: even if SuM_1 is compatible with SuM_2 and SuM_2 is compatible with SuM_3, it is not always the case that SuM_1 is compatible with SuM_3. As a consequence, in these comparisons all pairs of SuMs should be in principle taken into account, a constraint that can be hardly fulfilled in any practical situation. Therefore, an alternative strategy for obtaining intersubjectivity is generally pursued, based on the structural possibility of *calibrating M*.

In general, calibration is an operative procedure aimed at defining the symbolization function λ for each system *M*, so that even systems that cannot be guaranteed to be exact replications of each other produce in principle the same results for the same SuMs. In order to properly define a calibration process, the concept of *measurement standards* is to be introduced.

Measurement standards are instantiations of a given property to which given property values are conventionally assigned in order to be used as reference values. Measurement standards are then entities that, although outside the scope of RTM, are deemed essential component of any metrological system. Once a set of measurement standards is introduced, calibration is a process performed in two steps: determining a relation between the property values provided by measurement standards and selectable elements in the support *U* of the RS induced by the MS; using such information to determine a relation for obtaining a measurement result from each selectable element $u \in U$.[15,16]

Accordingly, the first step of the calibration can be formalized as a function mapping the set P of property values attributed to the available standards into the set U of selectable elements of the MS under calibration, $\gamma_{\mathrm{MS}}: P \rightarrow U, u = \gamma_{\mathrm{MS}}(p)$, being thus u the selectable element obtained from the measurement standard realizing the property of value p. In the simplest case it can be assumed that the mapping γ is both:

(i) injective, i.e., M has a resolution comparable with the adopted measurement standards, where resolution is to be intended as the smallest change in a property being measured that causes a change in the corresponding element of the support U, so that different standards produce distinguishable selectable elements of the MS;
(ii) surjective, i.e., measurement standards are available in the calibration stage for each selectable element in U, so that $\gamma(P) = U$.

Whenever these conditions hold, the mapping γ is invertible, and the application of γ^{-1} to any selectable element u obtained by the empirical application of M leads to a quantity value $\gamma^{-1}(u) \in P$. The second step of calibration specifies then that $V = P$, i.e., the set of possible measurement results coincides with the set of property values realized by the measurement standards exploited in calibration, and $\lambda = \gamma^{-1}$, i.e., the symbolization function is equal to the inverse of the calibration function. *Measurement is then obtained as a specific case of pre-measurement, for which the symbolic stage of the process is performed according to the information acquired by calibration.* As a consequence, a MS is characterized as a 4-tuple $\langle M, \mathrm{P1}, \mathrm{P2}, \mathrm{P3}' \rangle$, where M, P1 and P2 are defined as for PMSs, and P3′ specifies a symbolization function as obtained by inverting the calibration function γ.[17] Provided that suitable measurement standards

[15] The present definition is consistent with the definition of calibration provided by the VIM, according to which calibration is the "operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication" (JCGM 2008a). This definition is readily embedded in our context by interpreting the selectable elements of U as the indications the VIM refers to (and, of course, by neglecting measurement uncertainties, a subject that we are not explicitly dealing with here).

[16] The requirement that measurement standards are used as references operatively implies their widespread availability, so that in principle the same measurement standard should be able to interact with all measurement systems M, producing a selectable element, i.e., an indication, for each of them. The assurance that this condition holds is the main task for metrological systems: the techniques employed to this goal—the adoption of stable measurement standards that are traceable to even more stable measurement standards and so on in the traceability chain, up to a so-called primary measurement standard—are a basic subject of the literature of measurement science for physical quantities, and will not be further discussed here.

[17] A few extensions to this basic characterization can be immediately considered. In particular:
–the calibration function γ could be non-injective, in the (realistic) situation the system M has a lower resolution than the one of the set of the adopted measurement standards; in this case calibration induces a canonical partition P^* on the domain of γ and the symbolization function can be obtained by inverting γ^*: $P^* \longrightarrow U$;
–whenever measurement uncertainties are taken into account, a rule must be introduced to properly propagate them both in the calibration stage, i.e., while defining the function γ, and in the symbolization stage of measurement, i.e., while defining the function λ.

are available, the process can be performed on any MS, so that measurement results obtained by different MSs become comparable and thus intersubjective.

This justifies our basic claim that measurement structurally requires the proper (as specified by P1 and P2) usage of a calibrated (so that P3′ can be applied) MS. The stability, selectivity, and non-invasiveness of *M* and the stability of the measurement standards used for calibration are the crucial conditions for guaranteeing the fundamental features of measurement: objectivity and intersubjectivity.

4.1 Further remarks on Examples 4 and 5.

Let us keep into account once more the PMSs introduced in Examples 4 and 5 to shortly discuss how they can be calibrated and then exploited as actual MSs.[18]

In the case of Example 4 selectable elements are sets of reference objects. Let us assume, for the sake of simplicity, (i) that reference objects fulfill the condition of mutual equivalence, as specified in P1 of the user's manual, (ii) that calibration is performed by 1 g measurement standards (uncertainty issues are still out of scope here), and (iii) that, by operating on the balance, it is experimentally obtained that the two pans reach the same position when $2n$ standards are on one pan and n reference objects are on the other one. Given the stated hypotheses, the calibration function $\gamma(n\text{g}) = 0.5n$ ref is inferred, where n ref denotes the set of n reference objects. Finally, by inverting γ, the symbolization function λ such that $\lambda(n \text{ ref}) = 2n$ g is obtained.

In the case of Example 5 selectable elements are elongations of the dynamometer spring. Let us assume (i) that elongations are measured as lengths with a 1 mm resolution, (ii) that calibration is performed by set of standards realizing n g, $n = 1, \ldots 10$, and (iii) that, by operating on the balance, the calibration function $\gamma(n \text{ g}) = 5n$ mm is experimentally inferred. Thus, by inverting γ and assuming γ^{-1} can be linearly interpolated, the symbolization function λ such that $0.2n \text{ g} = \lambda(n \text{ mm})$ is introduced. Accordingly, the dynamometer spring behavior is described by the Hooke's law, $F = -kx$, being $k = 0.2$ g/mm (neglecting the contribution of gravitational acceleration) the spring constant of the given MS. Notice that this is a case of the important class of MSs whose interaction with SuMs is performed by a sensor, i.e., a transducer device sensitive to the measurand. The calibration function is defined here so to approximate the transduction function, indeed for dynamometers mapping weights to lengths. Since $\lambda = \gamma^{-1}$, this is a case in which such a function can be properly considered as operating a measurand reconstruction.

From this basis, two exemplary extensions can be introduced.

(1) The term k could be recognized as dependent on an influence quantity, e.g., temperature, $k = k(\text{T})$, so that the calibration function γ would hold only at the 'reference temperature', i.e., the temperature maintained when the MS is calibrated. In the case the temperature measured during the measurement process does not coincide with such reference temperature, a model must be adopted of the spring behavior dependence on temperature, formalized as a correction to the

[18] Property values in this section have been chosen to be expressive, not necessarily realistic.

calibration function γ itself and then to λ. Once more, the symbolization function operates a measurand reconstruction.

(2) The measurand whose value must be obtained could functionally depend on the sensor input quantity instead of being that quantity itself, e.g., the potential energy $E = 0.5kx^2 = 0.5Fx$ stored in the spring instead of the weight of the applied object. In this case, traditionally characterized as *indirect measurement*, the symbolization function plays the double role of reconstructing the input quantity and embedding such a model.

4.2 Further remarks on Example 6.

Once the distinction between pre-measurement and measurement is understood as grounded on calibration, it becomes clear why introducing a process of measurement in soft sciences is so difficult. According to RTM, that assumes measurement to be the construction of a morphism between an empirical and a numerical RS, the role of calibration is immaterial, so that no crucial difference is identified between measuring, for example, lengths or weights of bodies and competences or preferences of human beings. On the contrary, according to our model the possibility of calibrating MSs, and in particular the availability of measurement standards, is critical to discriminate between pre-measurement systems and measurement systems, and therefore to justify the claim of intersubjectivity for measurement results.

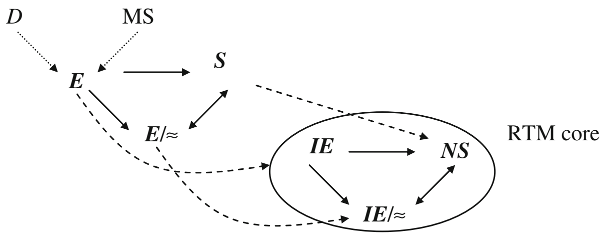
4.3 Comparison with RTM

The main features of the general model of measurement that has been introduced in this paper can be synthesized as follows.

(1) States of SuMs are modeled as elements of empirical RSs as the result of the very construction of MSs, not by an unspecified activity of idealization:
   (1.1) elements in the domain of the empirical RS are not necessarily empirically accessible;
   (1.2) relations in the empirical RS are empirically decidable by construction.
(2) It is not necessary to introduce any axioms on the relations in the empirical RS:
   (2.1) the problem of truth of such axioms does not arise;
   (2.2) the problem of checking of such axioms does not arise.
(3) RSs formalizing measurement results are not necessarily modeled as numerical RSs:
   (3.1) elements of such RSs are symbols, not necessarily numbers;
   (3.2) relations among elements of such RSs are empirically meaningful only if MSs support them.
(4) Measurement concerns SuMs, not empirical structures, and a measurement is given if and only if a MS is suitably exploited so that:
   (4.1) the objectivity of its results is guaranteed to a degree sufficient to a given aim;

(4.2) the intersubjectivity of its results is guaranteed to a degree sufficient to a given aim.

According to this model a measurement process is performed by accomplishing some, both empirical and theoretical, steps, as sketched in the following diagram.



(I) Empirical steps

Step 1: a MS is constructed, designed so to be able to interact with an intended domain of systems *D* with respect to the measurand under consideration.

Step 2: the MS is applied to one or more elements of the domain *D*, and the empirical RS ***E*** is determined by the specifying the relations induced by the MS into *D*.

Step 3: the empirical RS ***E*** is mapped to the quotient RS $\boldsymbol{E}/\approx$ obtained by classifying the elements that are not distinguishable by the MS. This mapping preserves the induced structure, because of the way the MS operates (in particular as specified in Part 2 of its user's manual).

Step 4: the quotient RS constructed in Step 3 is mapped to a symbolic RS ***S***.

(II) Theoretical steps

Step 5: the empirical RS is linked to an idealized empirical RS ***IE***, the idealization consisting in the assumption that the system satisfies a definite set of axioms characterizing the structural relations.

Step 6: these axioms and the possibility of performing given numerical operations on measurement results, as obtained by mapping ***IE*** to a numerical RS ***NS*** through the quotient RS $\boldsymbol{IE}/\approx$, are justified, at the idealized level, by proving existence and uniqueness theorems.

Within this general framework, the drawbacks that in this paper have been highlighted on RTM do not constitute a problem anymore.

(1) The problem of RTM abstractness is overcome at once, since the connection with the practice of measurement is embodied in the requirement that any measurement process must be performed by means of a MS.
(2) The problem that the RTM conditions of measurability are too strong to be necessary is met by distinguishing between concrete and ideal systems: the ideality of the axioms on the relations of an empirical RS is a consequence of the step from ***E*** to ***IE***, since ***IE*** is an idealization of an empirical RS.
(3) The problem that the RTM conditions of measurability are too weak to be sufficient is met by requiring that only properties of empirical SuMs a MS can interact with can be measured, so that "*no MS, no measurement*" is an analytical truth.

(4) The distinction between numbers and numerals as entities for expressing measurement results is grounded on the distinction between symbols in $\boldsymbol{S}$ interpretable as signs for numbers and symbols that do not allow such an interpretation.

Finally, it is worth noticing that the general model presented in this paper consists not only in the external addition of new components to the model of measurement provided by RTM. On the contrary, as idealizations of an actual empirical RS, any idealized empirical RS is tightly connected with a MS, since the relations in the RS are induced by the interaction with the MS. Accordingly, a question such as "Is it possible to measure a property of elements of D?" cannot be interpreted as "Is it possible to construct a homomorphism from an empirical RS $\boldsymbol{E}$ to a numerical RS $\boldsymbol{NS}$?", because $\boldsymbol{E}$ is not given unless a MS is applied to elements of D. Furthermore, a question such as "Is it possible to measure a property of objects in $\boldsymbol{E}$?" cannot be interpreted as "Is it possible to construct a homomorphism from $\boldsymbol{E}$ to $\boldsymbol{NS}$?", but is to be intended as "Is it possible to find out an idealized RS $\boldsymbol{IE}$ such that $\boldsymbol{IE}$ is both (i) a satisfactory model of $\boldsymbol{E}$ and (ii) a system embeddable in $\boldsymbol{NS}$?". According to this general model, the first part of the last question asks for a justification of the fact that some axioms are introduced as adequate to describe the empirical relations in $\boldsymbol{E}$, and thus for a justification of the capability of the MS to induce such relations.

5 Conclusion

According to the general model we have proposed, the basic issues to address in order to understand a given measurement process are (i) what are the intended SuMs; (ii) what are the MSs. Furthermore, a solution to issue (i) is implicit in the solution of (ii), inasmuch the set of SuMs, for a given MS, could be directly defined as the set of systems the MS is able to interact with. Since in RTM such issues find no place, we conclude that RTM, although significant, is too abstract to offer a satisfactory account of what a measurement process is.

To summarize, the presented model is characterized by the following statements.

1. Measurement is a process requiring the empirical interaction of a system under measurement (SuM) and a measurement system (MS).
2. A MS is such that the result of this interaction is an empirically selectable (traditionally, but not necessarily: observable) element of a given set (actually: a relational system, i.e., a set equipped with relations).
3. The measurand value obtained as measurement result is given by a mapping (actually: a morphism) from the set of the empirically selectable elements (point 2) to a set of symbols.
4. Such a mapping at least embeds the information obtained by calibrating the MS as follows:
   4.1. the MS is put in interaction with one or more measurement standards (according to point 1, as if standards were peculiar SuMs), each of them realizing a measurand value that is assumed to be known, and the empirically selectable elements obtained as the result of the interaction (according to point 2) are recorded and used to infer a calibration function;

   4.2. the mapping (point 3) is obtained by inverting the calibration function (point 4.1) in the way previously described.
5. Points 1–4 are basic specifications that can be variously extended by further characterizing the formal mapping (point 3). In particular:
   5.1. The empirical interaction performed by the MS (point 1) could be recognized depending on one or more influence quantities, and calibration could be performed under environmental conditions different with respect to those of measurement. In this case, if a model of the MS behavior dependence on such influence quantities is available, the mapping (point 3) could embed the information derived from this model, so to correct the effects / the distortion generated by influence quantities.
   5.2. The measurand could be different with respect to the quantity according to which the MS interact with the SuM, and functionally depending on it. In this case, the mapping (point 3) could embed such a function, so to actually generate information on the measurand.

Hence, points 1–4 characterize a model that is general enough to include as particular cases those described in point 5. Furthermore, this model is able to describe 'direct' measurement (point 4), measurement with corrections keeping into account effects of influence quantities (point 5.1), and 'indirect' measurement (point 5.2). Finally, the model is general enough to describe measurements performed according to different measurement methods, for both physical and 'soft' systems, and capable to give an account of the differences between physical and 'soft' measurement, by also highlighting the usual lack of definite calibration procedures when MSs for 'soft' measurement are introduced.

Acknowledgements We are grateful to two anonymous referees for their helpful comments on an earlier draft of this paper. The third author gratefully acknowledges the importance of Ludwik Finkelstein's work and continuous support for the development of many concepts underlying this paper.

References

Birkhoff, G. (1948). Lattice theory (American Mathematical Society colloquium series, Vol. 25).
Bridgman, P. (1927). *The logic of modern physics*. New York: Macmillan.
Carnap, R. (1966). *Philosophical foundations of physics*. New York–London: Basic Books, Inc.
Finkelstein, L. (2003). Widely, strongly and weakly defined measurement. *Measurement, 34*, 39–48. doi:10.1016/S0263-2241(03)00018-6.
Finkelstein, L. (2005). Problems of measurement in soft systems. *Measurement, 38*, 267–274. doi:10.1016/j.measurement.2005.09.002.
Hempel, C. G. (1952). *Fundamentals in concepts formation in physical sciences*. Chicago: University of Chicago Press.
Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. Glencoe: Free Press.
JCGM. (2008a). JCGM 200:2008, International vocabulary of metrology—basic and general concepts and associated terms (VIM, 3rd ed.). Downloadable from http://www.bipm.org.
JCGM. (2008b). JCGM 100:2008, Evaluation of measurement data—guide to the expression of uncertainty in measurement (GUM, 1995 with minor corrections). Downloadable from http://www.bipm.org.
Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. I). New York: Academic Press.

Kuhn, T. S. (1961). The function of measurement in modern physical sciences. *Isis, 52*, 161–193. doi:10.1086/349468.

IEC. (2008). IEC 60050 series, and Electropedia (also known as the "International Electrotechnical Vocabulary (IEV) Online"). Retrieved from http://www.electropedia.org.

Luce, R. D. Krantz. D. H., Suppes, P. & Tversky, A. (1990). *Foundations of measurement* (Vol. III). San Diego: Academic Press.

Mari, L. (2003). Epistemology of measurement. *Measurement, 34*, 17–30. doi:10.1016/S0263-2241(03)00016-2.

Mari, L. (2007). Measurability. In M. Boumans (Ed.), *Measurement in economics* (pp. 41-77). Amsterdam, Elsevier.

Narens, L. (1985). *Abstract measurement theory*. Cambridge: Mass MIT Press.

Narens, L. (2002). *Theories of meaningfulness*. Mahwah, NJ: Lawrence Erlbaum Associates.

Niederée, R. (1992). What do numbers measure? A new approach to fundamental measurement. *Mathematical Social Sciences, 24,* 237–276. doi:10.1016/0165-4896(92)90063-B.

Pfanzagl, J. (1968). *Theory of measurement* (2nd ed.). New York: Wiley (Vienna: Physica, 1971).

Roberts, J. (1979). *Measurement theory*. Reading, Mass: Addison-Wesley.

Scott, D. & Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic, 23*, 13–128. doi:10.2307/2964453.

Stevens, S. (1946). On the theory of scales of measurement. *Science, 103*, 677–680. doi:10.1126/science.103.2684.677.

Suppes, P. (2006). Transitive indistinguishability and approximate measurement with standard finite ratio-scale representations. *Journal of Mathematical Psychology, 50*, 329–336.

Suppes, P., Krantz. D. H., Luce, R. D., & Tversky, A. (1990). *Foundations of measurement* (Vol. II). New York: Academic Press.

Suppes, P., & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.). *Handbook of mathematical psychology* (Vol. 1, pp. 1–76). New York, Wiley.