# Epistemological Foundations of Measurement

**Luca MARI**
**Università Cattaneo - LIUC**
**21053 / Castellanza (VA), Italy**

## ABSTRACT

Some notes for a conceptual history of measurement are presented, to introduce the basic positions that have been maintained about the issue that can be considered as the general epistemological problem of measurement: what characterizes measurement with respect to generic evaluation?
This characterization, in which three distinct positions are recognized correspondingly to what we call Metaphysical Period, Anti-Metaphysical Period, and Relativistic Period, allows us to introduce and shortly discuss some general issues on the current epistemological status of Measurement Science.
The goal of this work is more to identify and characterize some current epistemological problems of measurement than to give them a solution.

**Keywords:** Foundations of Measurement, Measurement Theory, Measurement Science, Measurement and Evaluation, Measurement Uncertainty

## 1. INTRODUCTION

In the Plato's Theaetetus, Socrates explains the meaning of a principle attributed to Protagoras with the following words: «things are to you such as they appear to you, and to me such as they appear to me». The sophistic principle that Socrates was commenting on is the one, well-known, according to which «man is the measure of all things, of the existence of things that are, and of the non-existence of things that are not». In his analysis, Socrates considers that Protagoras was asserting the equivalence of *knowledge* and *sensation*: things can be known because they can be perceived, and they are known as they are perceived. It is interesting that this position, according to which «the same thing that *appears* warm to me and cold to you *is* warm to me and cold to you», establishes a paradigm of knowledge in reference to the concept of *measure*.

According to this meaning, measurement is simply synonym of evaluation, with the consequence that also estimations, personal judgments, and possibly even random assignments should be considered as specific examples of measurements. The issue here is only partly definitional (we could also be uninterested in definitions, although they sometimes maintain a useful role in guaranteeing some chance of mutual understanding). The fact is that this subjectivistic standpoint seems to be completely unable to justify the peculiar epistemic status usually recognized for measurement: Physics is, or was, paradigm for all the sciences mainly because of its ability

to objectively measure quantities, and then define quantitative relations among them.

Measurement is a specific kind of evaluation, i.e., it is an operation aimed at associating an *information entity*, the result of measurement, with the *state of the system* under measurement in reference to a given quantity, the measurand.

Even if not unconditionally adhering to the praises that can be easily located in the scientific literature about the role of measurement (e.g., «the progress of civilization is in strict relation to the development of measures» [1]: it is clear that nothing similar would be stated about generic evaluations), it is plausible that measurement has something special with respect to a generic evaluation. We believe that the quest for this peculiarity can be aptly regarded as *the general epistemological problem of measurement*:

Problem A:
> *provided that measurement is an evaluation,*
> *what characterizes measurement*
> *with respect to generic evaluation?*

It should be clear that this *problem* relates more to the issue of «what is it useful for?» than of «what is it?», and therefore its answers have operative, and not only terminological, implications.

In the course of history the epistemic peculiarities of measurement have been found in at least three distinct areas:
- *ontological reasons* (measurement is an evaluation able to determine those numbers that are essential properties of things);
- *formal reasons* (measurement is an evaluation producing symbols that can be formally dealt with in a well definite way);
- *informational reasons* (measurement is an evaluation whose results are informationally adequate to given goals).

Goal of the present work is to sketch some basics for a critical analysis of these three interpretations, and then to investigate some consequences of such an epistemological reconstruction for the current status of Measurement Science.

## 2. SOME ETYMOLOGY

The term «measure» comes from the Latin *mensura*, abstract form derived from *mensus*, past participle of the

verb *metiri*, in its turn derived from the ancient *mitis*, present with minimal variations not only in Latin and in Greek, but also in the German and Indian areas. This term plausibly meant *wisdom, measure in psychological sense*. Some traces of this acceptation remain today in several idiomatic forms, e.g., «beyond measure», *exceeding some given correct limit*, and «within measure», *with moderation* (analogously, the Italian idiom «*usare due pesi e due misure*», literally «to use two weights and two measures» means *to be partial and thus unfair*). The Greek verb μετρειν (*metrein*) conveys the concept of measurement as evaluation, judgment, as witnessed by the corresponding noun μητισ (*metis*), good judgment, wisdom, prudence. In the Hesiod's Theogony the goddess Μητισ, *wise advice*, is introduced as the mother of Pallas (Athena), known to be the most cultured among gods and human beings, able to freely assume various shapes, and «the maker of all right things»: this therefore, according to Greek mythology, was the daughter of the «goddess of measure»; these qualities result from measurement.
According to this meaning, measurement is then essentially a *wise, subjective evaluation*.

Etymology allows us to identify a second class of meanings for the concept of measurement. By extending the root of the verbs *metiri* and μετρειν we obtain *me-s* and *me-ns*, the month (the Italian term, *mese*, clearly maintains this derivation) thought of as a unit to measure periods of time on the basis of the lunar cycle (the term *moon*, and the German *mond*, keep trace of this origin).
This acceptation, thus oriented to connote the measure as an *objective description*, was deeply influenced by the Euclidean model of geometry. According to Euclid (Elements, Book V, definitions 1-3), «a magnitude is a part of a magnitude, the less of the greater, when it measures the greater; the greater is a multiple of the less when it is measured by the less; a ratio is a sort of relation in respect of size between two magnitudes of the same kind», on the hypothesis that a (geometrical) quantity is measurable because it can be expressed as the ratio of due (integer) numbers.

Some quantities were found to be not compliant with such a definition, as the case of the diagonal of a square whose side is the unit: the use of «incommensurable» (in-com-*mens*-surable) with the meaning of *impossible to compare by measure* (and therefore *irrational*, because *impossible to express as a ratio*), *lacking a common measure*, recalls this paradigmatic role of the Euclidean geometry.

But until recently the emergence of such «non-measurable» quantities was not sufficient to break up the monolithic framework built around the Euclid's definition, and its only consequence has been a spread of terms to denote the same concept of *quantity under evaluation by empirical means*: «magnitude», «observable», «parameter», «state variable», «property», «attribute», …, and particularly «dimension» (di-*mens*-ion…), sometimes even adopted as synonym of *something measurable* (as it is plausibly the usage in the case of «Dimensional Analysis»).
This terminological anarchy has led to some curious idiomatic expressions, such as «measuring and counting» (as in the title of the seminal paper by H. Helmoltz, 1887

[2]) or «weights and measures» (as in «International Committee for Weights and Measures»), as if counting or weighing were not specific techniques to accomplish a measurement.

Problem B:
> *is measurement a subjective evaluation*
> *or an objective description?*

## 3. SKETCHES FOR A CONCEPTUAL HISTORY OF MEASUREMENT: THE METAPHYSICAL PERIOD

«When you can measure what you are speaking about, and express it in numbers, you know something about it; otherwise your knowledge is of a meager and unsatisfactory kind». This statement by Lord Kelvin, 1889, is often quoted to report what we could now consider the traditional standpoint on Problem B, and therefore implicitly on Problem A (the fact that we deem this position as "traditional" does not mean that we consider it just a matter of the past: on the contrary, it is plausible that in the current teaching such a position is still the most widely adopted when introducing some epistemology of measurement).
According to this position, measurement *derives its objectivity from a property of the measured things*: «each thing that is accessible to our knowledge has a number, since without numbers we can neither understand nor know» (from an excerpt of Pythagorean school). And in analogous terms: «the elements of numbers were supposed to be the elements of all things, and the whole heaven a musical scale and a number» (Aristotle, Metaphysics).
The formulation given to this hypothesis by Galileo Galilei (Il Saggiatore, 1632) is well-known: the «great book of nature» cannot be understood «but by learning its language and knowing the characters in which it is written: it is written in mathematical terms».
Therefore:

Position 1:
> *Measures are inherent properties*
> *of the measured things*

(and more properly, in reference to the Euclidean standpoint: measures are properties of the relation, and specifically of that particular kind of relation that is the ratio, between measured things and the chosen measurement unit.)

This Position reached a widespread consensus, well beyond the scientific community. For example, the aim of obtaining objectivity has been recognized as a main reason for the introduction of the metric system during the French revolution: «the meter, by "de-humanizing" the measures and making them independent of man, "objective" to him and morally neutral, transformed them from an instrument of the human arrogance to a means able to ease the comprehension and the collaboration among people» [3].

Indeed, in spite of the hypothesis that «numbers are in the world» (as Kepler wrote in his Letter to Micheal

Maestlin, 1595), measures have been used in the past to give a social justification to some arbitrariness: «what would the destruction of the feudal system be useful for if the Lords will remain free to increase or decrease their measures as much as they wish?» (so declared a popular protest in 1791, reflecting the expectations that followed the first period of the French revolution; quoted in the Kula's book).

On the other hand, it is usual to deal with numbers related to things and conveying information not only on things themselves but on the more complex system constituted by the things and their environment, thus possibly including some *human factors*.

It is the typical case of the prices to which things are bought and sold: how could they be considered objective, i.e., related *only* to their object and independent of its context?

We believe that this problem – the assumption that «numbers are in the world», together with the recognition that at least some of them cannot be simply assumed as objective – was the reason that led to introduce the traditional distinction between «primary» and «secondary qualities»: only the former would claim to be objective, and therefore «in the proper sense» measurable, properties.

(just to quote the classical Campbell's book [4]: «the basis of the Locke's between primary qualities, independent of observation, and secondary qualities, dependent on observation, is simply the distinction between qualities which, if they are observed at all, will always be observed to be the same and those which may appear different to different observers».)

The successes of the experimental method during the XVII and XVIII centuries laid new bases to make the empirical activity of measurement characteristic of the scientist work. Trace of the objectivism that was deemed essential to this activity can be recognized in the conceptual foundation of the Theory of Error proposed by K. F. Gauss at the beginning of the XIX century: any physical quantity has its own *true value*, so that the experimental variability of the measurement results is explained as deriving from the introduction of *errors*: «while analyzing the meaning of the measures that he had obtained, the experimenter tries to guess the true value, the value that would have produced the best achievable instrument» [5].

Residual of this standpoint is plausibly the definition of «true value» that is given by the *International vocabulary of basic and general terms in metrology* (the «VIM», issued by ISO in 1993 [6]): «the value which characterizes a quantity perfectly defined, in the conditions which exist when that quantity is considered», while noting immediately that «the true value of a quantity is an ideal concept and, in general, cannot be known exactly».

A rather curious definition for a concept that should ground an empirical activity…

## 4. SKETCHES FOR A CONCEPTUAL HISTORY OF MEASUREMENT: THE ANTI-METAPHYSICAL PERIOD

Measurement is so fundamental for science that it should not be amazing that changes in Philosophy of Science directly reflect on the interpretation of measurement, and therefore on the answers given to what we have previously called the «general problem of measurement», i.e., our Problems A and B.

In the explicit attempt to remove all the metaphysical assumptions from the scientific knowledge, in the course of the first decades of the XX century the neopositivistic school tried to make the very concept of verification clear, by grounding it on sensation, observation, and – ultimately – measurement.

«Phenomena do not contain anything of numerical, but only our sensation. We can introduce numerical concepts by establishing a procedure to measure them. Numbers are assigned to the nature by ourselves, because phenomena exhibit only the qualities we observe», as R. Carnap wrote in 1966 [7].

Measurement was then becoming «protocol of truth», the privileged means to establish the ultimate criterion to decide about the truth or falsehood of statements. In this perspective, empirical sciences could defer the responsibility to decide on the degree of truth of their models and leave it to measurement, as witnessed by the epistemic significance assigned to the classical concept of «crucial experiment».

It is interesting that, in spite of its anti-metaphysical claims, NeoPositivism maintained in fact the term «true value», and put it at the core of its implied concept of *observable pure data*, although trying to keep it behind the scenes and sometimes making it somehow fuzzier, as in the concept of «conventional true value», a manifest oxymoron.

(emblematic of this troubled situation is the shift from «*the* true value» to «*a* true value», as can be retrieved from (relatively) recent standard documents; the former is, e.g., in [8], while the latter, rather hard to understand – how can something actually true be at the same time indeterminate? –, is in the already quoted [6].)

With such a definitely non-empirical basis, the need arose to replace the (claimed) criterion of verification with *a* (conceptually weaker but empirically useful) *criterion of acceptance*, in order to support the decision on the quality of measurement, and therefore, again, on the specificity of measurement as a particular evaluation.

Since an *empirical* reference *external* to the measurement procedure was lacking, the opposite target of a *formal* reference *internal* to the procedure was looked for.

The mentioned paper by H. Helmoltz [2] was the basis for the work that led S. Stevens to formulate a Theory of Scales [9], according to which «measurement is the assignment of numerals to objects or events according to rule, any rule» [10]. From an epistemological point of view, the important contribution of such a Theory is related to its analysis of the distinctive factors for the different scale types. Peculiar of, e.g., ratio vs. interval scale type is the set of relations that are preserved by

scale transformations, and therefore the algebraic *structure* in which the scale values are embedded.

The fact is that such a structure is not an inherent characteristic of the measured quantity, but depends on the *state of knowledge* on that quantity (for example, temperature was deemed to be an interval quantity before its thermodynamic re-definition, that introduced a unique, fixed point in the scale, thus making it a ratio quantity). Assuming that scale types are structured as a (partially) ordered set (so that, e.g., interval «is less than» ratio type, because scale transformations in the first case are algebraically weaker than in the second case), with the corresponding relation «greater than» thought of as «conveys more information than» [11], we can conclude that in foundations of measurement *ontology was being replaced by epistemology*.

Moreover, while the available knowledge on a quantity implies an upper bound on its scale type, the specific procedure adopted to evaluate a quantity determines the scale type in which the obtained values are actually embedded (it is not hard, for example, to show how to evaluate lengths by means of a procedure producing only ordinal values, as in the case of a sieve adopted as comparator). With this further characterization, the classical categorization of *concepts* as qualitative or quantitative (or, in a somehow more refined way, classificatory, ordinal, and quantitative, as in [7] and [12]), becomes a categorization of *operative procedures*: *epistemology was being replaced by technology.*

From our point of view, these changes are even more relevant when considering that the Theory of Scales (and its generalization, usually designated as Measurement Theory *par excellence*, that formalizes measurement as a morphic mapping from an empirical to a symbolic relational system [13] [14]), does not imply any specific empirical criterion to characterize measurement, whose definition is only based on the compliance with a formal condition.

Position 2:

*Measures are results of operations*
*that preserve the relations observed*
*among measured things*

In these terms the claimed equivalence measurement = morphic mapping, while identifying a necessary condition for the *correctness* of an evaluation (F. Roberts introduced the concept of *meaningfulness* at this regard; some objections to this point of view can be found in [15]), can be hardly considered sufficient to grasp the peculiarity of measurement as an objective operation. (indeed in [14] F. Roberts presents several examples of completely subjective evaluations, such as estimations of degrees of preference among alternatives, and calls them measurements whenever the condition of morphism holds for them; the already quoted statement by S. Stevens, «measurement is the assignment of numerals to objects or events according to rule, any rule» is clearly in this line, being «the rule» something that can be formalized as a morphism.)

The extreme consequence of this so-called «representational point of view to a Theory of Measurement» (but actually become *the orthodox* Measurement Theory) is the idea that measurement can be characterized in purely mathematical terms, as the concept of Abstract Measurement Theory [16] implies.

## 5. SKETCHES FOR A CONCEPTUAL HISTORY OF MEASUREMENT: THE RELATIVISTIC PERIOD

Quantum mechanics made the existence of a «measurement problem» clear (where / what is the truth in the microscopic world?), but until recently the assumption was maintained that in macroscopic realms measurement is entitled to reach truth.

A new philosophical position (the fundamental book by T. Kuhn, 1969 [17] is sometimes taken as the turning point; the P. Feyerabend's work, 1975, [18] can be quoted as the one which brought such a position to its extreme consequences) completed the anti-metaphysical process initiated by Neopositivism, now largely just historical matter.

In our perspective it is of great interest to begin by re-analyzing the distinction, considered fundamental by the neopositivistic school, between the so-called *theoretical* and *observational* terms. While the former are recognized to acquire a meaning only in the context of a model, the latter would refer to *directly observable* objects, properties, or events, thus playing the role of actual pillars of any system of empirical knowledge. An observational term has been described as «a descriptive (nonlogical) term which may occur in a quickly decidable sentence, a "quickly decidable sentence" being defined in its turn as a singular, non-analytic sentence such that a reliable, reasonably sophisticated language user can very quickly decide whether to assert it or deny it when he is reporting on an occurrent situation» [19].

In the current cultural climate, characterized by a widespread epistemic relativism, this «definition» sounds really fuzzy and has lost much of its plausibility: *raw sense data simply do not exist*. Even the terms that are traditionally considered observational, such as those appearing in the statements that express measurement results, are currently recognized «theory-laden» (according to Hanson's terminology).

This conceptual shift has an immediate implication in terms of a (usually deemed) fundamental distinction in measurement, namely the one between *direct and indirect* (or derived) measurement. According to the approach adopted by the recent ISO *Guide of expression of Uncertainty in Measurement* (GUM) [20], the expression of any measurement result requires the indication of a measurand value and an estimation of its uncertainty, and both of them depend for their evaluation on the previous measurement of a given set of recognized influence quantities. Being each of these influence quantities a new measurand, the mentioned dependence should be recursively applied: this would imply that no measurement could be ever completed! This recursive process is then usually arrested by assuming that the influence quantities are not in themselves influenced by other quantities, thus on the hypothesis that they can be «directly measured». On the other hand, it is clear that

15

from a theoretical point of view this has to be considered an approximation.

The transition is then complete: *once thought of as epistemologically fundamental, observational terms are now recognized as the expression of the operative need to make any empirical process of information acquisition approximate.*

As Bridgman wrote [21], «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. […] By forcing the physical experience into the straight jacket of mathematics, with its yes-no sharpness, one is discarding an essential aspect of all physical experience and to that extent renouncing the possibility of exactly reproducing that experience. In this sense, the commitment of physics to the use of mathematics itself constitutes, paradoxically, a renunciation of the possibility of rigor».

With the very concept of true value disappearing as simply meaningless, *no external and absolute references remain* to evaluate the quality of measurement results and to distinguish measurement from generic (morphic) evaluation. Properly speaking, *errors* become only a very peculiar case of a far more general parameter, always required when expressing a measurement result: *uncertainty.*

Let us quote from the mentioned GUM: no formal technique for evaluating the measurement uncertainty can be the «substitute for critical thinking, intellectual honesty, and professional skill; [...] the quality and utility of the uncertainty quoted for the result of a measurement ultimately depends on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value».

Is then measurement, properly speaking, just a *methodology*? Methodologies are useful, but their basic requirement is that *they should work*, surely not their truth…

Could we assert, for example, that the procedure presented by the GUM to evaluate and formally express the uncertainty in measurement is «truer» than the alternative methods it is aimed at supplanting? The term of comparison is surely in this case *adequacy*, not truth.

Position 3:
> *Measures are results of operations*
> *recognized as adequate for their goal*
> *of obtaining information on measured things*

We have intentionally stated this latter Position in very generic terms. Indeed, our tentative reconstruction of a «conceptual history of measurement» comes to an end here.

If an answer to *the general epistemological problem of measurement* cannot be metaphysical in nature (as for Position 1) and a formal characterization is not enough for it (as for Position 2), we believe that the search for the

reasons of the commonly recognized «special adequacy» of measurement (as for Position 3) is the correct way to cope with such a problem.

## 6. SOME FURTHER COMMENTS (WHY THE TRANSITION WAS UNAVOIDABLE)

The shift *from error to uncertainty* is far more than purely terminological: on the contrary, it is paradigmatic of a modification in the epistemology of measurement.

True values, but also sense data obtained by direct observations, have a kind of *Platonic reality*, perhaps useful in modeling phenomena (for the same reason why many models of physical systems are based on differential calculus even though no measuring system has infinite resolution) but unmaintainable from an operative point of view.

Numbers are symbols (i.e., information entities) for (components of states of) empirical entities: numbers do not belong to the physical world.

Considering the following:
Statement 1:
> «at the instant of the measurement
> the thing is in a definite state»

Statement 2:
> «at the instant of the measurement
> the measurand has a definite value»

While traditionally such statements would be plausibly considered as synonymous, their conceptual distinction is a fundamental fact of metrology: the former represents a basic assumption of measurement (neglecting here the issues related to the role of the measurement in quantum mechanics); the latter is epistemologically meaningless (and – however – operationally irrelevant). Measurement results are informational, and not empirical, entities: what in measurement is *determined*, and therefore considered pre-existing, is the thing state, and not the measurand value, that is instead *assigned* on the basis of the instrument reading and the calibration information [22].

To acknowledge that *measurement is not a determination* (as in «measurement is the set of operations having the object of determining the value of a quantity» [6]) *but an assignment* (as in «measurement is the process of empirical, objective assignment of numbers to the attributes of objects and events of the real world, in such a way as to describe them» [23]) is the most explicit signal of a transition from the «classical» paradigm [24].

With the assumption that measurement is an assignment, it is recognized that any measurement result reports information that is meaningful *only in the context of a metrological model*, such a model being required to include a specification for all the entities that explicitly or implicitly appear in the expression of the measurement result: the thing under measurement as identified with respect to its environment (manifesting its interdependence with the thing through the presence of some influence quantities); the reference standard adopted for calibration and the implied traceability chain to the primary standard; the procedure used to perform

16

the measurement; finally, the symbol(s) formalizing the measurand value and its estimated uncertainty.

The complexity of this model directly influences the objectivity of the obtained measurement results: but even in the best attainable case a non-null «intrinsic uncertainty» remains [25] [20], as a trace of the inherent differences between «the realm of things we say» and «the realm of things we do» (already quoted from [21]).

(the extreme example of intrinsic uncertainty is for counting, i.e., measurement performed in absolute scale, in which no admissible scale transformations are allowed. In this case the standard coincides with the unit of measurement, that is operatively defined by the criterion to decide whether to include a «candidate» thing in the counting or to discard it. The implementation of such a criterion could take all possible cases into account only if provided with an infinite amount of information. On the other hand, the simplicity of this class of metrological models usually makes them the best option for many measurement problems, in which measurands are actually evaluated in terms of related countable quantities. It is known that these measurands, such as time durations when measured by counting a number of discrete periodic events, are measured with the lowest uncertainties.)

It is important to point out that what we have called Position 3, that can be characterized as a relativistic, if not even subjectivistic, and information-dependent standpoint, does not prevent from including some objective components in its models. A rational subjectivistic judgment exploits all the available sources of information, and selects them according to their quality, coming from either *objective* or *subjective* sources (in this perspective, we believe that subjectivistic positions could be considered more general than objectivistic ones: but this is another topic).

This complementarity of objectivism and subjectivism in measurement was substantially recognized in a fundamental recommendation issued in 1981 by the *International Committee for Weights and Measures* (CIPM) [26] and integrally adopted by the GUM. According to such a recommendation, uncertainty of measurement results can be evaluated by means of two distinct and complementary methods:
- some uncertainties, conventionally designated as «of type A», are computed as suitable statistics of experimental data, usually obtained as repeated instrument readings;
- some other uncertainties, conventionally designated as «of type B», are instead estimated on the basis of the observer's personal experience and the available a priori information, and therefore express a *degree of belief* on the possible measurand values.

Hence the evaluation of uncertainties in measurement could, and usually should, require the formalization *of beliefs*. In these terms, the transition is complete: *ontology and the criterion of truth have been replaced by information and a criterion of adequacy*.
But this epistemological position makes what we have called the general epistemological problem of measurement even more critical: is measurement then just an evaluation whose results are subjectively believed adequate to given goals?

The fact that the great majority of the people operatively involved in physical measurement would not embrace such a position can be a convincing stimulus to motivate the search for more cogent and specific answers to the problem, but it is not in itself a proof of the falsehood of the position.

# 7. ON THE EPISTEMOLOGICAL STATUS OF (A) MEASUREMENT SCIENCE

Complementary to the problems we have discussed up to now, some investigation can be made on that systematically organized body of knowledge commonly called Measurement Science (MS) [27]. The issue is not related, obviously, to the existence of such a body of knowledge, but to its actual *nature of science*:

Question1:
is MS to be aptly considered
a science (empirical or formal),
or just a technology / a methodology?

MS could be, and actually is, the aggregate of distinct components; but if a MS properly exists, it should be more than the gathering of elements taken from other sciences, such as Physics (e.g., laws on transduction effects) and Mathematics (e.g., signal theory).
The fact is that the body of knowledge termed MS constitutes a rather anomalous discipline. It is hardly reduced to a formal science, but it does not seem to be characterized in terms of its contents (although contents are usually adopted to internally partition its scope, as in the case of Electrical Measurement, Mechanical Measurement, Thermal Measurement, …): in some sense, MS stands between Physics, Mathematics, and Engineering. This hybrid nature is the origin of some fundamental issues on the epistemological status of MS. According to our current conception of the role of measurement:

Question2:
can a MS, properly speaking, exist?

and, in derivative way:

Question 3:
does a MS exist today?

It should be clear that we are looking for an *epistemological* answer to such questions: to state that «by MS it is meant what the researchers involved in measurement do» is nominalistic, and therefore immaterial to this case. Moreover, since the existence of a Technology of Measurement is not under discussion, our questions could be reformulated as:

Question 4:
is what the researchers involved in measurement do
integrally technology,
or can at least a part of it
be properly considered (empirical or formal) science?

We are not interested in «definitions»: our issue concerns the basic criteria that should be adopted in the validation of the results:

* an empirical science is expected to produce results that, at least in principle, are passible of experimental falsification;

* a formal science is expected to produce results consistent to a given set of axioms;

* a technology is expected to produce results complying with given goals

(and, but this is obvious, all should produce results claimed to be somehow useful).

With a few diagrams let us try to introduce and comment on these alternatives.

An *empirical science* can be thought of as an organized body of knowledge describing some aspects of the physical world and allowing predictions on them:
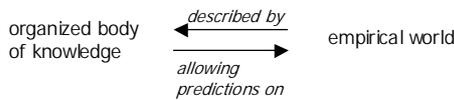


The concept of a *formal science* is defined in terms of a suitably formalized language, called a *formal system*, provided with axioms and inference rules allowing to derive expressions of the language (the theorems) from the chosen axioms, the set of language expression thus identified being called a *theory*:
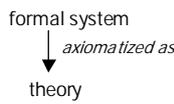


While a formal system is a purely syntactical entity, a *model* for a formal system is a structured set aimed at interpreting the formal system in reference to its elements.

In the case that an empirical science is formalized, a combined version of the previous two diagrams then applies:
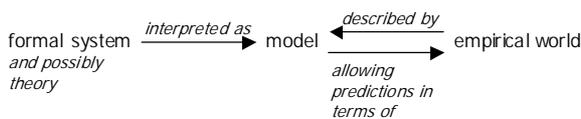


How can measurement be interpreted in terms of these diagrams? In other terms: is what is commonly called MS to be aptly considered an empirical science or a formal science (or whatever else)?

Although a clear-cut answer is not easily found in the current scientific literature, we suggest that a solution to this problem can be reconstructed from the axiomatization of measurement that D. Krantz, R. Luce, P. Suppes, and A. Tversky proposed in [13], in the context of the representational point of view. Given the hypothesis that measurement can be formalized as a morphic mapping from an empirical to a symbolic relational system, they assert that both these relational systems are models of the same theory, their (common) scale type:
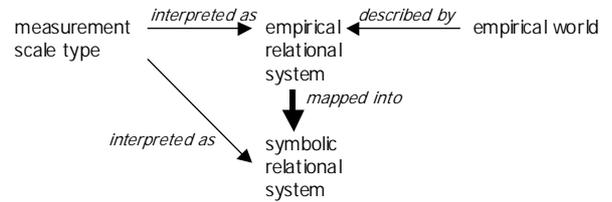


the bold arrow thus being the representation of measurement.

This explanation is surely quite extreme, since no arrows are drawn from the model(s) back to the empirical world, so that MS would be in principle unfalsifiable: the representational point of view gave an important contribution, but only as far as the consistency internal to the modeling structures (both the empirical and the symbolic relational systems) is concerned.

Let us comment on this point with an example, as basically taken from [14]. A subject is asked to express its preference among the elements of a given set, say musicians; it is then suggested that a ranking evaluation *is* a measurement whenever the condition of ordering among symbols is consistent with the «empirical» comparison, i.e., the mapping associating symbols to musicians is a morphism. Our conclusion: by means of such an operation some information has been obtained on the evaluating subject, and therefore on the model, but surely not on the elements of the set, i.e. the empirical world!

We consider a really unfortunate fact that people interested in (empirical…) Measurement Science are still forced to rely on these theoretical bases only, that do not take into account fundamental topics for actual measurement such as measuring instruments and standards, calibration and traceability.

## 8. NEXT STEPS

«To insist on calling these other processes [the quotation refers here to non-physical methods of measurement] measurement adds nothing to their actual significance, but merely debases the coinage of verbal intercourse. 'Measurement' is not a term with some mysterious inherent meaning, part of which may be overlooked by the physicists and may be in course of discovery by psychologists. It is merely a word conventionally employed to denote certain ideas. To use it to denote other ideas does not broaden its meaning but destroys it: we cease to know what is to be understood by the term when we encounter it; our pockets have been picked of a useful coin» [28].

This *ante litteram* objection to the representational point of view was aimed at defending the «classical» meaning of measurement against its semantic extensions proposed by psychologists: it was the attempt to provide an answer to the general epistemological problem of measurement by simply denying the very existence of the problem.

But this argument is contradictory: were the meaning of «measurement» only conventional, a suitable agreement would be sufficient for its extension. We believe instead that «measurement» *has* an inherent meaning: the characterization of measurement with respect to generic evaluation must be based on the identification of such a

meaning, that a suitable epistemological analysis should be able to progressively reveal (and then make less and less «mysterious»).

The direction we suggest to follow is operational: what is measurement for?

While the answer «to obtain and formally express information on a system» is too generic (it would hold more or less for *any* evaluation), the specification «to obtain and formally express empirical and objective information on a system» could be correct but does not make such criteria of «being empirical and objective» explicit.

The pragmatic aim of evaluation is to enable symbolic processing of data drawn from the empirical world, so that any result obtained in data processing can be re-interpreted in terms of the measured things:
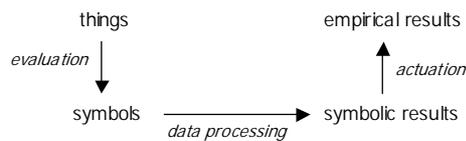


Crucial for the validity of this re-interpretation is therefore the *faithfulness* of the operation that associates symbols with things:
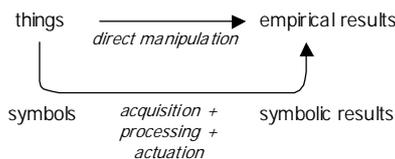


so that the *same* empirical results can be obtained by direct manipulation and the more complex but usually more efficient sequence of operations: acquisition, processing, and actuation.

Such a faithfulness is the ultimate criterion of adequacy for an evaluation.

A radically operational position could now declare *any evaluation recognized as adequate in this sense to be a measurement*. This position has several epistemological implications, and among them:

- the status of measurement for an evaluation is not a yes-no characteristic: the more is an evaluation recognized as faithful, the more it can be dealt with as a measurement; therefore the status of measurement for an evaluation is not given a priori but it is a target to be reached;
- the decision whether an evaluation should be considered a measurement does not depend solely on formal or empirical characteristics, and it must take into account also some pragmatic components;
- a measurement is aimed at producing results adequate to given goals, and not «ideal» or «true» values; the declaration of their estimated faithfulness is thus integral part of the expression of the measurement results.

In order to understand the commonly recognized acceptation of (physical) measurement, the condition of adequacy is clearly not sufficient. We suggest that it must be complemented by two further requirements:

- the results of a measurement should convey the same information to different observers, i.e., they should be

inter-subjectively communicable, so that different observers should obtain the same results by processing the same data by the same procedure;
- the results of a measurement should convey information related only to the measured thing, and not to its environment, there including the observer, so that the same thing should produce the same results when measured in different conditions and by different observers.

For the sake of synthesis, these requirements can be expressed as:

measurements are
intersubjective and objective evaluations

This characterization is able to justify the *role of measuring systems* (the role basing the rough, but ultimately correct, position according to which *measurements are evaluations performed by means of measuring systems*), aimed at assuring intersubjectivity by means of the adequate choice of properly traceable standards and objectivity by means of the adequate setup and usage of instruments whose selectivity make them sensitive to the measurand and insensitive to all other quantities [29].

Therefore measuring systems are not only the objects of operative measurement: first of all, *they are the essential component for making us able to give consistent epistemological foundations to measurement*. We are firmly convinced that any future Measurement Theory will be grounded on the formalization of their structure and their operative function in that specific kind of evaluation that is measurement.

# 9. REFERENCES

[1] A. Pérard, *Les mesures physiques*, Paris: Presses Universitaires de France, 1947.

[2] H. Helmoltz, *Zählen und Messen*, 1883.

[3] W. Kula, *Measures and men*, Princeton: Princeton University Press, 1986.

[4] N. R. Campbell, *Physics: the elements*, Cambridge: Cambridge University Press, 1920.

[5] J. Idrac, *Measure et instrument de measure*, Paris: Dunod, 1960.

[6] ISO, *International vocabulary of basic and general terms in metrology*, Geneve: ISO, 2nd ed., 1993.

[7] R. Carnap, *Philosophical Foundations of Physics*, New York: Basic Books, 1966.

[8] DIN, *Basic concepts in metrology. Terminology relating to the uncertainty of measurement and the assessment of measuring instruments and measuring equipment*, DIN 1319 - Part 3, 1983.

[9] S. Stevens, "On the theory of scales of measurement", *Science*, Vol. 103, No. 2684, 1946, pp. 677-680.

[10] S. Stevens, "Measurement, psychophysics, and utility", in *Measurement: definitions and theories*, C. West Churchman and P. Ratoosh (eds.), New York: Wiley, 1959, pp. 18-63.

[11] L. Mari, "Notes towards a qualitative analysis of information in measurement results", *Measurement*, Vol. 25, No. 3, 1999, pp. 183-192.

[12] C. G. Hempel, *Fundamentals of concept formation in empirical science*, Chicago: The University of Chicago Press, 1952.

[13] D. Krantz, R. Luce, P. Suppes, A. Tversky, *Foundations of measurement*, New York: Academic Press, Vol.1 (1971), Vol.2 (1989), Vol.3 (1990).

[14] F. Roberts, *Measurement theory – With applications to decision making, utility, aand the social sciences*, Reading: Addison Wesley, 1979.

[15] E. Adams, "On the nature and purpose of measurement", *Synthese*, Vol. 16, 1966, pp. 125-169.

[16] L. Narens, *Abstract measurement theory*, Cambridge: MIT Press, 1985.

[17] T. S. Kuhn, *The structure of scientific revolutions*, Chicago: University of Chicago Press, 1969.

[18] P. Feyerabend, *Against method: outline of an anarchistic theory of knowledge*, Atlantic Highlands: Humanities Press, 1975.

[19] G. Maxwell, "The ontological status of theoretical entities", in *Scientific knowledge: basic issues in the philosophy of science*, J. A. Kourany (ed.), Belmont: Wadsworth, 1987, pp. 348-354.

[20] ISO *et al.*, *Guide to the expression of uncertainty in measurement*, Geneve: ISO, 1993.

[21] P. W. Bridgman, "How much rigor is possible in physics?", in *The axiomatic method with special reference to geometry and physics*. Proceedings of an international symposium held at the University of California, Berkeley, December 26, 1957-January 4, 1958, L. Henkin, P. Suppes, A. Tarski (eds.), Amsterdam: North-Holland, 1959, pp. 225-237.

[22] L. Mari, G. Zingales, "Uncertainty in Measurement Science", in *Measurement Science – A discussion*, K. Karija, L. Finkelstein (eds.), Tokyo: Ohmsha, 2000, pp. 135-147.

[23] L. Finkelstein, "Measurement: fundamental principles", in *Concise encyclopedia of measurement and instrumentation*, L. Finkelstein, K. Grattan (eds.), Oxford: Pergamon, 1994, pp. 201-205.

[24] L. Mari, "The role of determination and assignment in measurement", *Measurement*, Vol. 21, No. 3, 1997, pp. 79-90.

[25] UNI, *Misure e misurazioni. Termini e definizioni fondamentali*, Italian Standard UNI 4546, 1984 [in Italian].

[26] CIPM, *BIPM Proc.-Verb. Com. Int. Poids et Mesures*, Vol. 49, No. 26, 1981, pp. 8-9 (in French).

[27] L. Finkelstein, "Measurement science", in *Concise encyclopedia of measurement and instrumentation*, L. Finkelstein, K. Grattan (eds.), Oxford: Pergamon, 1994, pp. 205-209.

[28] J. Guild, "Are sensation intensities measurable?", in *Interim report of the Committee appointed to consider and report upon the possibility of quantitative estimates of sensory events* (Part III), A. Ferguson (Chairman), Report of the 108th annual meeting of the British Association for the Advancement of Science, Cambridge, 1938.

[29] L.Mari, "Beyond the representational viewpoint: a new formalization of measurement", *Measurement*, Vol. 27, No. 2, 2000, pp. 71-84.