

AN INTERDISCIPLINARY PROBABILISTIC THEORY OF MEASUREMENT

Giovanni Battista Rossi
DIMEC - Università degli Studi di Genova,
Via All'Opera Pia 15 a, I-16145, Genova, Italy
gb.rossi@dimec.unige.it

Abstract

Measurement plays a fundamental role, in both physical and behavioural sciences, as a link between abstract models and empirical reality. A challenge therefore arises: is it possible to develop a unique theory of measurement for the different domains of science it is involved in? A basic kernel for one such theory, dealing with two main issues, the measurement scale and the measurement process, is presented. Since uncertainty plays an essential part, the theory is expressed in probabilistic terms. Some notes on the application of this theory to psychophysical measurements are also provided.

Although measurement has been fundamental to the development of modern science since the time of Galileo and Newton, in-depth investigation into its establishment, involving both physicists and psychologists, came relatively late, in the second half of the nineteenth century. So it is no wonder that a major contribution came from the eclectic genius of Helmholtz (1887). He proposed a fruitful analogy between measuring and counting: when the characteristic we wish to measure *may be thought of* as the amount of something, then measurement is similar to counting the number of elementary units of that something. Interestingly enough, Fechner (1860) had also invoked the paradigm of counting for measuring the intensity of a sensation. In this case, zero is the perception threshold and the units are the elementary increments in the sensation evoked by barely perceptible variations in the corresponding stimulus. Despite this promising beginning, development of the common understanding of measurement encountered a crisis at the beginning of the twentieth century with the Report of the British Association for the Advancement of Science (Ferguson et al., 1940). A Committee appointed to discuss and report on the possibility of a “quantitative estimation of sensory events” concluded with deep divisions between two parties holding opposing views, thus generating negative consequences that have continued up to the present. Since then, remarkable achievements in measurement science have been attained by both parties, though communication and interaction has been scarce. Now there seem to be good opportunities for a reconciliation: in this regard, an excellent environment has been provided by the European Call on “Measuring the Impossible” and by the related Coordination Action MINET (Measuring the Impossible Network), chaired by Birgitta Berglund. In this framework, inter- and multi-disciplinary projects have been launched and several opportunities for discussion, dialogue and constructive interaction have been provided. One of the major challenges at present is the development of a common theory of measurement (Rossi & Berglund, 2009).

Basic questions

What do we expect from an interdisciplinary theory of measurement?

Three main questions need to be addressed, perhaps, namely,

- what is the meaning of measurement,
- how do we measure and
- what can be measured.

In this paper, a basic kernel is presented for one such theory, organised into two main parts, the measurement scale and the measurement process, dealing with the first two questions respectively. An answer to the third question, concerning measurability, derives from the combination of the two main parts of the theory, rather than remaining a separate issue. First, each of the two main parts are presented in deterministic terms, then the theory is reformulated in probabilistic terms in order to account for *uncertainty* as an inherent feature of measurement. Concerning language, what we (want to) measure will be called a *characteristic* and what manifests such a characteristic will be called an *object*, irrespective of whether it is an object, an event or even a person. A summary of the proposed theory is presented in Table 1, which will be discussed throughout the paper.

Table 1. Synopsis of the proposed theory.

The measurement scale		
Scale type	Deterministic approach	Probabilistic approach
Order	$a \succ b \Leftrightarrow m(a) \geq m(b)$	$P(a \succ b) = P(x_a \geq x_b)$
Interval	$\Delta_{ab} \succ \Delta_{cd} \Leftrightarrow$ $m(a) - m(b) \geq m(c) - m(d)$	$P(\Delta_{ab} \succ \Delta_{cd}) =$ $= P(x_a - x_b \geq x_c - x_d)$
Ratio	either $a \square b \circ c \Leftrightarrow m(a) = m(b) + m(c)$ or $a/b \succ c/d \Leftrightarrow \frac{m(a)}{m(b)} \geq \frac{m(c)}{m(d)}$	either $P(a \square b \circ c) = P(x_a = x_b + x_c)$ or $P(a/b \succ c/d) = P\left(\frac{x_a}{x_b} \geq \frac{x_c}{x_d}\right)$
The measurement process		
Process	Deterministic approach	Probabilistic approach
Observation	$y = f(x)$	$P(y x)$
Restitution	$\hat{x} = f^{-1}(y)$	$P(x y) = P(y x) \left[\sum_{x \in X} P(y x) \right]^{-1}$; $\hat{x} = \mu(x y)$
Measurement	$\hat{x} = h(x) = x$	$P(\hat{x} \mu) = \sum_{y \in Y} \delta[\hat{x} - (x y)] P(y x)$

The measurement scale

Deterministic Framework

The most accredited way of clarifying the meaning of measurement is given by the representational theory (Krantz et al., 1971-1990) in which the notion of *measurement scale* is central. In its general meaning, a scale includes the conditions that enable measurement, that is, it identifies the set of objects that manifest the characteristic under consideration, specifies its empirical properties and the indicates the existence of object to number mapping so as to reproduce empirical relations in a numerical domain. This led to the identification of a few basic scales types that are presented in Table 1. As this theory is assumed to be generally well-known, we shall make no attempt to survey it. We shall simply point out that each scale is characterised by a *representation theorem*, that shows how empirical relations are mapped into the corresponding numerical ones, and by a *uniqueness theorem*, concerning the class of admissible transformations. In the top-left part of the Table we have considered three types of

scales (ordinal, interval and ratio) and we have indicated the corresponding representation theorems. For ratio scales, we have considered two possible alternative representations: one based on an empirical operation of *addition*, denoted by the symbol “ \circ ”, and the other based on an empirical relation of *ratio*, denoted by the symbol “ $/$ ” (Miyamoto, 1983). The former representation is typical of physical measurement and refers to empirical *extensive* structures, whilst the latter is more appropriate for measuring the intensity of a sensation. We therefore propose to call the corresponding empirical structure *intensive* (Rossi & Crenna, 2009).

Why a probabilistic approach is needed

If interpreted in a deterministic way, the empirical relation $a \succ b$ implies that whenever we compare objects a and b , we always observe that a is greater than or equal to b . Clearly this, in general, is too strict a requirement: if the objects are intuitively close to each other, we may expect to sometimes observe a greater than or equal to b , but other times a less than b . This more general scenario may be properly described in probabilistic terms by assigning $P(a \succ b)$, $P(a \square b)$ and $P(a \prec b)$, where “ P ” denotes the probability of such observations occurring.

Probabilistic representations

Consider, for example, the deterministic representation for weak order,

$$a \succ b \Leftrightarrow m(a) \geq m(b). \quad (1)$$

By adopting a probabilistic logic, $a \succ b$ is no longer a statement that may be either true or false, rather it expresses a *possibility* having a certain degree of probability, $P(a \succ b)$. Consequently, the assignment of numbers to objects will no longer be unique. Rather, each element, say a , will be associated with a *random variable*, x_a (Regenwetter, 1996). The representation now implies that the probability of observing a relation between two objects is the same as the probability of obtaining the same kind of relation between the associated random variables. So, in the case of order, the representation theorem becomes

$$P(a \succ b) = P(x_a \geq x_b). \quad (2)$$

The representation theorems for the three scale types under consideration are indicated in Table 1 (top-right part) (Rossi, 2006). These probabilistic representations are useful, in physical measurements, for providing a basis for expressing the uncertainty of primary standards, as well as for addressing the evaluation of key comparisons. In the case of psychophysics, they may be useful for evaluating and expressing the uncertainty of the results of scaling experiments (BIPM, 2008).

The measurement process

The Measuring System

Once we have constructed a reference scale, we then have to consider how to measure objects that are not included in the scale. This requires a *measurement process* which, in physical measurement, usually implies the use of a *measuring system*. Since we wish to develop an

interdisciplinary theory, we must look for a general enough definition of a measuring system that can also be applied to psychophysical measurement. We thus propose to define it as an empirical system capable of interacting with objects that manifest the characteristic under investigation and, as a result of such interaction, capable of producing *signs* that can be used to assign a measurement value to the measurand (the characteristic to be measured), in agreement with a previously defined reference scale. We suggest that this definition may be used not only in physics and engineering, where the measuring system or instrument is usually a physical device, but also in psychophysics, where people act as measuring instruments, or even in psychometrics, where the “instrument” is a procedure based on test items (Rossi & Berglund, 2009). Then, looking for a general model of the measurement process, we propose to parse it in two sub-processes, *observation* and *restitution*, as illustrated in Figure 1.

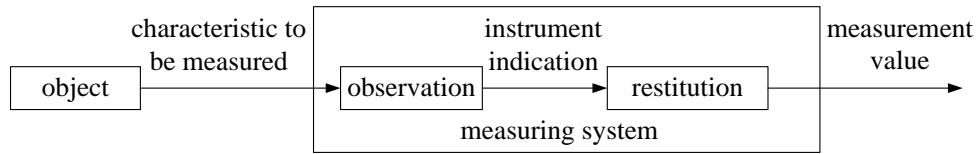


Figure 1. Measurement process scheme

In the observation phase, the measurand is inputted to the measuring system that produces a sign or indication. In the restitution phase, this indication is *interpreted* on the basis of the *calibration (or observation) function*, that is, a function which characterises the behaviour of the measuring system, usually obtained by calibration, and the measurement value is obtained. Measurement may be understood as the concatenation of observation and restitution. Let us now translate this qualitative description into a simple mathematical model. If we call the calibration function f , *observation* may be described by

$$y = f(x), \quad (3)$$

restitution by its inversion,

$$\hat{x} = f^{-1}(y), \quad (4)$$

and *measurement* by the concatenation of observation and restitution,

$$\hat{x} = f^{-1}[f(y)] = x, \quad (5)$$

which ideally reduces to a unitary transformation. This is a deterministic model, describing an ideal measurement, but it is unable to properly represent what happens in reality. We now therefore have to reformulate it in probabilistic terms.

The Probabilistic Model

Let us first reconsider observation. In contrast with the ideal model, when we input to the measuring system a measurand whose value is x , and we repeat the experiment several times, in general we do not always obtain the same indication, y , but rather a cluster of indications. An appropriate way to describe this behaviour is to assign a probabilistic distribution to the indications, conditioned on the value of the measurand, as shown in Figure 2a. The

observation function $y = f(x)$ is now replaced by the conditional probability distribution $P(y|x)$. For example, Figure 2a shows that if the input value is $x = 6$, three indications may be obtained, 10, 12 and 14, with conditional probabilities of $P(y=10|x=6) = 0.2$, $P(y=12|x=6) = 0.6$, $P(y=14|x=6) = 0.2$.

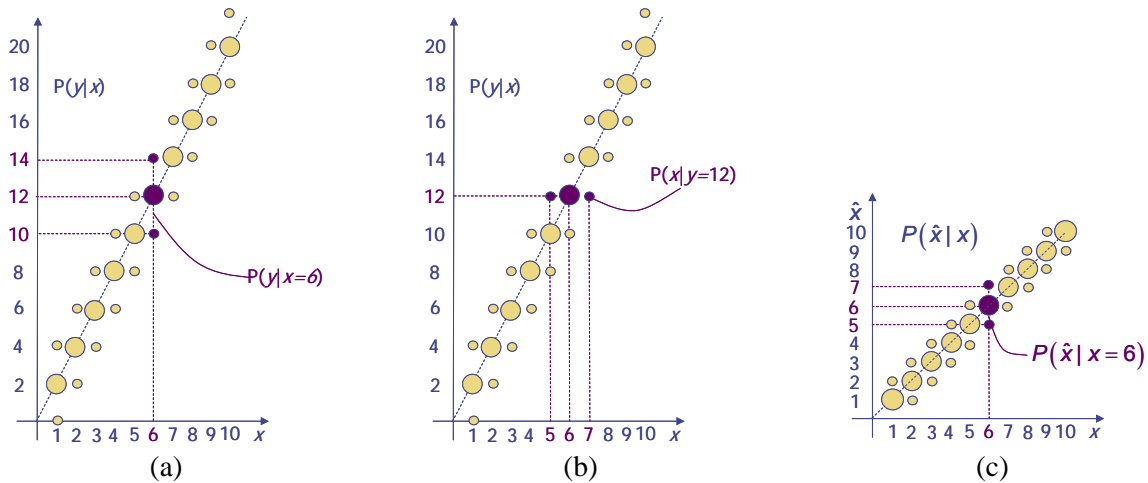


Figure 2. A probabilistic model of the measurement process

How can we achieve restitution? The idea is very similar to the deterministic case. Suppose we observe $y = 12$: Figure 2b shows that such an indication may have been *caused* by three values of x , 5, 6 and 7, and from the same graph we may obtain the probabilities of these three causes, $P(x=5|y=12) = 0.2$, $P(x=6|y=12) = 0.6$ and $P(x=7|y=12) = 0.2$. What we have just done is a *probabilistic inversion* of the observation transformation, that may be analytically obtained by applying the well-known Bayes-Laplace rule (Press, 1989). Lastly, we may combine observation and restitution in order to obtain a description of the overall measurement process, provided by the distribution $P(\hat{x}|x)$ (Figure 2c). A summary of these formulae is presented in the bottom-right part of Table 1 (Rossi, 2006). We think that this description of the measurement process is appropriate for physical measurements, where the measuring system may be identified in a physical instrument, but we suggest that it may also be used to describe measurement in psychophysics, where people act as measuring systems, as happens, for example, with the *master scaling* method by Berglund (1991).

Measurability

Once the meaning of measurement has been clarified by the notion of measurement scale and the way we measure by the concepts of measuring system and measurement process, we may conclude by briefly mentioning the issue of measurability (Finkelstein, 2005; Rossi, 2007). As there is no room here for a thorough discussion of this topic, we simply propose, on the basis of what has been discussed so far, a simple and straightforward measurability criterion. We suggest that that *a characteristic x of a class of objects is measurable if the following four-step procedure can be successfully applied:*

1. define the *class of objects* that manifest the characteristic;
2. identify the *empirical properties* that define the characteristic;
3. construct a *reference measurement scale*;
4. devise at least one *measuring system* based on that reference scale.

It is hoped that this definition can be used as a shared basis for addressing measurability, in both physical and psychological sciences.

Final remarks

We have presented a basic kernel for a probabilistic theory of measurement, addressing three key questions, namely, what is the meaning of measurement, how do we measure and what can be measured. We have suggested that the answer to the first question can be found in a theory of the measurement scale, to the second in a general model of the measurement process and to the third in a conceptually simple measurability criterion. Further study of many aspects is required, including the problem of derived scales, a deeper characterisation of the measurement process and the extension to multi-dimensional measurement, to name just a few. In developing this theory, we have carefully avoided linking it to special classes of measurements or to specific measuring techniques or technologies. We therefore hope that it can act as a good starting point for achieving a common view of measurement among different scientific disciplines.

References

- BIPM (2008). Evaluation of measurement data – Guide to the expression of uncertainty in measurement. (JCGM 100:2008).
- Berglund B. (1991). Quality assurance in environmental psychophysics. In S.J. Bolanowsky & G.A. Gescheider (Eds.), *Ratio Scaling of Psychological Magnitudes – In Honor of the Memory of S.S. Stevens*. Hillsdale: Erlbaum; pp.140-162.
- Fechner, G. T. (1860). *Elemente der Psychophysik*. Volume I. English trans.: New York: Holt, Rinehart and Winston, 1966.
- Ferguson, A., Myers C.S., & Bartlett R.J. (1940). Quantitative estimation of sensory events - Final BAAS Report. *Advances of Science*, 2, 331-349.
- Finkelstein, L. (2005). Problems of measurement in soft systems. *Measurement*, 38, 267-274
- Helmholtz, H. (1887). Zählen und Messen Erkenntnis – theoretisch betrachtet, *Philosophische Aufsätze Eduard Zeller gewidmet*. Fuess, Leipzig.
- Krantz D. H., Luce, R.D., Suppes, P., & Tversky A. (1971-1990). *Foundations of Measurement*, Vol. 1-3. New York: Academic Press.
- Miyamoto, J. M. (1983). An axiomatization of the ratio/difference representation, *Journal of Mathematical Psychology*, 27, 439-455
- Press, S.J. (1989). *Bayesian statistics*. New York: Wiley.
- Regenwetter, M. (1996). Random utility representations of finite m-ary relations. *Journal of Mathematical Psychology*, 40, 219-234.
- Rossi, G. B. (2006). A probabilistic theory of measurement. *Measurement*, 39, 34-50
- Rossi, G. B. (2007). Measurability. *Measurement*, 40, 545-562.
- Rossi G. B., Berglund B. (2009). Measurement related to human perception and interpretation – State of the art and challenges. XIX IMEKO World Congress, Sept. 6-11, Lisbon.
- Rossi G. B., Crenna F. (2009). Ratio scales revisited. MINET Conference on Measurement, sensation and cognition, 10-12 November 2009, NPL, London, UK.