Measurement and Metrology Requirements for Empirical Studies in Software Engineering

Alain Abran, Asma Sellami

Draft August 26, 2002

Position paper proposed for the workshop

A Roadmap for Empirical Software Engineering

At STEP 2002, Montreal, Canada

Abstract

'Software metrics' are most often proposed as the measurement tools of choice in empirical studies in software engineering, and the field of 'software metrics' is most often discussed from the perspective referred to as 'measurement theory'. However, in other disciplines, it is the domain of knowledge referred to as 'metrology' that is the foundation for the development and use of measurement instruments and measurement processes. In this paper, our initial modeling of the sets of measurement concepts documented in the ISO International Vocabulary of Basic and General Terms in Metrology is used to investigate and position the measurement concepts referred to in the Guide to the Software Engineering Body of Knowledge. This structured analysis reveals that much work remains to be done to introduce the full set of measurement and metrology concepts as fundamental tools for empirical studies in software engineering.

1. Introduction

Over recent decades, hundreds of 'software metrics' have been proposed by researchers and practitioners alike, in both theoretical and empirical studies, for measuring software products and software processes [1,2,3,4,5,6,8,9]. Most of these metrics have been designed based either on the intuition of the researchers or on an empirical basis, or both. In their analysis of some of these metrics, researchers have most often used the concepts of 'measurement theory' as the foundation for their analytical investigation. However, while relevant, 'measurement theory' deals with only a subset of the classical set of concepts of measurement; 'software metrics' researchers, by focusing solely on 'measurement theory', have investigated mainly the representation conditions, the mathematical properties of the manipulation of numbers and the proper conditions for such manipulations [8,9]. Our survey of the literature on software metrics has not, however, come up with references to the classical concepts of metrology in these investigations into the quality of the metrics proposed to the software engineering community.

Section 2 presents our initial modeling of the set of concepts in this ISO Vocabulary. Section 3 presents a wider measurement process model and positions the metrology concepts model within this larger model. In section 4, the 2001 Trial Version of the IEEE and ISO 19759 Guide to the Software Engineering Body of Knowledge (SWEBOK) [13] is analyzed and the measurement-related topics identified; it is then positioned with respect to the subset of metrology concepts dealt with. Finally, recommendations for improving the foundations of software measurement tools for empirical studies in software engineering are presented in Section 5.

2. High-level model of the ISO vocabulary on metrology

2.1. The ISO Metrology Vocabulary

In empirical studies, including those in engineering as well as in other fields such as business administration and a significant number of the social sciences, measurement is one of a number of analytical tools. Measurement in these other sciences is based on a large body of knowledge; such a body of knowledge, built up over centuries and millennia, is commonly referred to as the field of 'metrology'. This domain is supported by government metrology agencies, which are to be found in most industrially advanced countries.

Quantitative methods for empirical studies in software engineering are most often based on 'software metrics'. To investigate how these software metrics map to the classic domain of metrology, we use the set of concepts contained in the ISO document that represents the official national and international consensus on the vocabulary of basic and general terms on metrology [7]. This ISO Vocabulary follows some of the concepts of the traditional presentation of vocabularies, with 120 terms described individually in textual descriptions. However, this mode of representation is challenging in terms of assembling the full set of interrelated terms; to improve the presentation and the understanding of this complex set of interrelated concepts, we presented in [11,12] an initial set of models for the various levels of metrology concepts within the ISO Vocabulary.

The high-level model of the set of categories of terms is presented in Figure 1. This model, together with some sub-models presented later on, correspond to our current understanding of the topology integrated into the vocabulary of this specialized area of the body of knowledge relating to metrology. To represent the relationships across the terms, the classical representation of a production process was selected: e.g. input, output and control variables, as well as the process itself inside the box. In Figure 1, the output is represented by the 'measurement results' and the process itself by the ‘measurement’ in the sense of measurement operations, while the control variables are the 'etalons' and the 'quantities and units'. This set of concepts represents the 'measuring instrument'. It is to be noted that the measurement operations, and, of course, the measurement results, are influenced by the 'characteristics' of the measuring instruments.

In the Vocabulary, the term 'measurements' used as a single term corresponds to the 'set of operations' used for measuring; this translates into the French 'mesurage'. Also, in all figures and tables in this paper, a term taken directly from the ISO Vocabulary will be appear in roman type, while terms representing concepts not specifically listed will appear in *italics*; for instance, in Figure 1, we have added the term 'Input', which is not included in any of the six categories of the ISO Vocabulary. Models of each of these six categories of metrology terms are presented next.

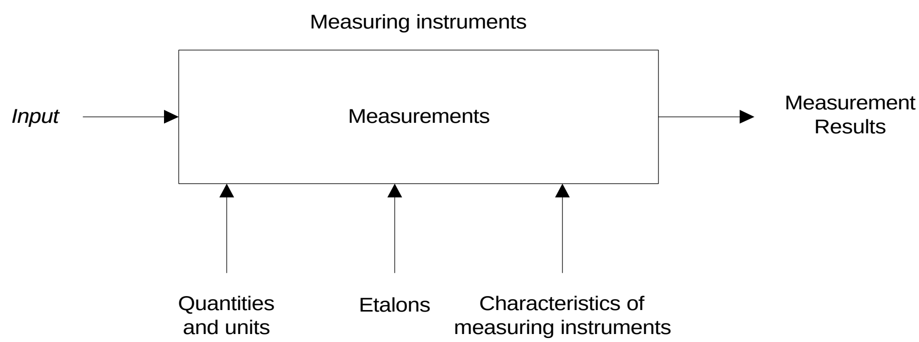


Figure 1: Model of the categories of metrology terms

2.2. The measurement foundation and measurement process

The term 'metrology' (Figure 2) includes all aspects of measurement (theoretical and practical), referred to collectively in the metrology literature as the science of measurement. Metrology encompasses the 'principles of measurement', which represent the scientific basis for measurement. From the principles of measurement, the 'method of measurement' in the general sense is then instantiated by a measurement as a set of operations. Figure 2 depicts this hierarchy of concepts.

The detailed topology of the measurement process is instantiated next in a 'measurement procedure' (Figure 3), again as a process model having several inputs, many control variables and an output representing the 'results of measurement'.

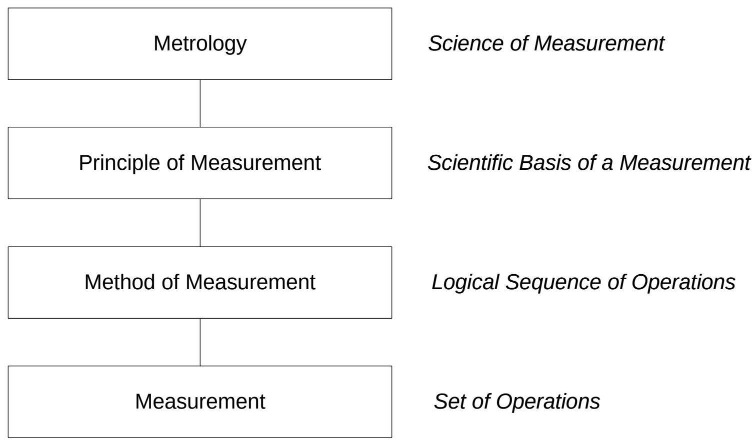


Figure 2: Measurement foundations

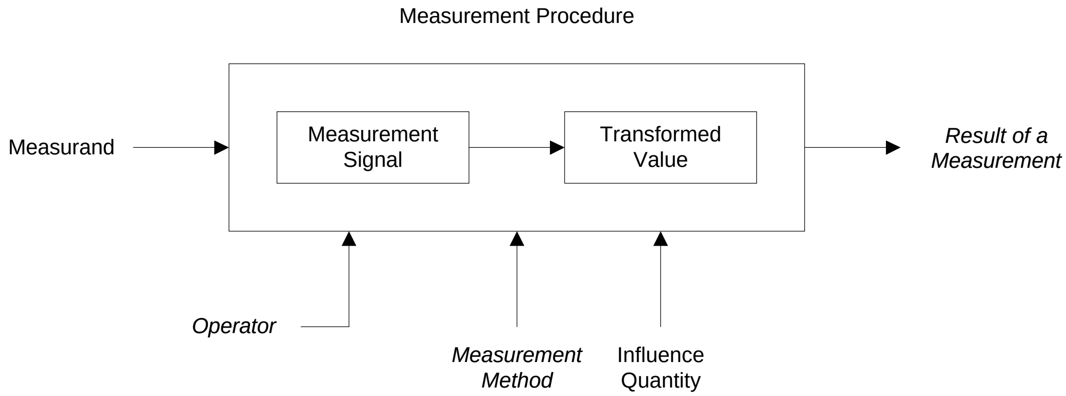


Figure 3: Measurement procedure

To carry out a measurement exercise, an operator should design and follow a 'measurement procedure' which consists of a set of operations, specifically described, for the performance of a particular measurement according to a given measurement method. The instantiation of a measurement procedure handles a 'measurement signal' and produces a transformed value which represents a given measurand. The results of the measurement can have been influenced by an 'influence quantity' during the measurement process: for example, the temperature of a micrometer during the measurement of the length of a particular object.

2.3. The Measurement Results

The category 'measurement results' is presented next in the form of a structured table according to the types of measurement results, the modes of verification of the measurement results and information about the uncertainty of measurement – Table 1. Again, this structure is our own.

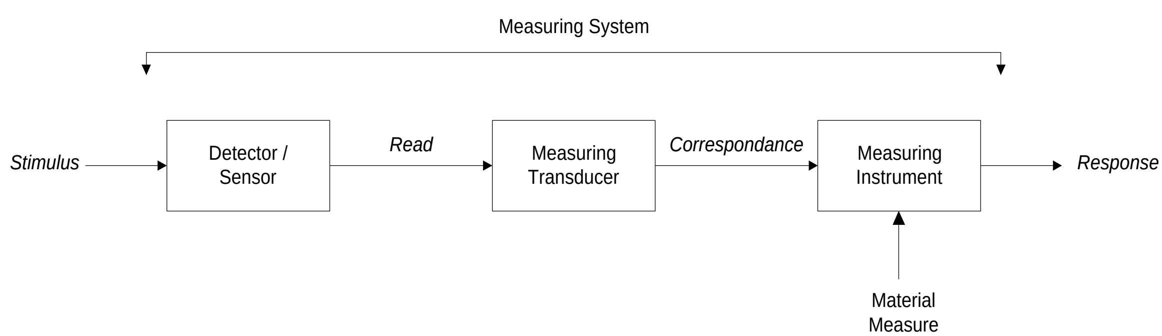
Table 1: Classification of terms in the category of 'Measurement Results' in [ISO 1993]

Types of measurement results	***Modes of verification of measurement results***	**Uncertainty of measurement**
Indication (of a measuring instrument) Uncorrected result Corrected result	Accuracy of measurement Repeatability (of results of measurements) Reproducibility (of results of measurements)	Experimental standard deviation Error (of measurement) Deviation Relative error Random error Systematic error Correction Correction factor

2.4. Measuring Instruments

Figure 4 presents a more detailed view of a measuring system, including a measurement instrument: following a stimulus or an input signal, the detector (or sensor) will detect the presence of the signal. If there is a signal, the instrument will indicate a value of a quantity associated with it. Following a reading, the measuring transducer provides an output quantity having a determined relationship to the input quantity. More detailed models are presented in the Appendix.

Figure 4. Model of a measuring system



2.5. Characteristics of measuring instruments

Table 2 presents our classification of the category of metrology terms addressing the 'characteristics' of the measuring instruments. This table is structured according to the quantitative and qualitative characteristics of a measuring instrument, the results of functionality tests to be performed at the moment of the utilisation and control of the instrument, and its measuring range.

Each of these characteristics will have an impact on the qualities of the measurement results, and on the quality of the models using these measurement results as their inputs.

Table 2: Classification of terms of 'Characteristics of Measuring Instruments' in [ISO 1993]

Quantitative	***Qualitative***	***Functionality test***		**Measuring Range / Working Range**
		Use	***Control***	
Rated operating conditions Limiting conditions Reference conditions Instrument constant Response characteristic Sensitivity Discrimination (threshold) Resolution (of a displaying device) Dead band	Stability Transparency Drift Response time Accuracy of a measuring instrument Accuracy class (*class index*) Freedom from bias (of a measuring instrument) Repeatability (of a measuring instrument)	Error (of indication) of a measuring instrument Maximum permissible errors / Limits of permissible error Bias (of a measuring instrument) Fiducial error (of a measuring instrument)	Datum error (of a measuring instrument) Zero error (of a measuring instrument) Intrinsic error (of a measuring instrument)	Nominal Range Span Nominal Value

3. Measurement process in Abran and Jacquet

In their work as ISO editors for the Guide to the Verification of Functional Size Measurement Methods (ISO 14143-3) [14], Abran and Jacquet studied the various authors dealing with 'metrics validation' [10,15,17]. They found significant variations in the authors' approaches as well as the use of similar terms by these authors, but with very significant differences in the related concepts. To clarify the confusion due to the inconsistent terminology used by these authors, Abran & Jacquet proposed a broader measurement process model (Figure 5). This model identifies 4 distinct steps, from the design of a measurement method to the exploitation of the measurement results [10]. Then, they positioned the approaches of the various authors, as well as appropriately positioning the validation concepts that were being addressed differently by these authors, depending on whether or not they were addressing validation issues related to Steps 1 to 4 of the process model in Figure 5.

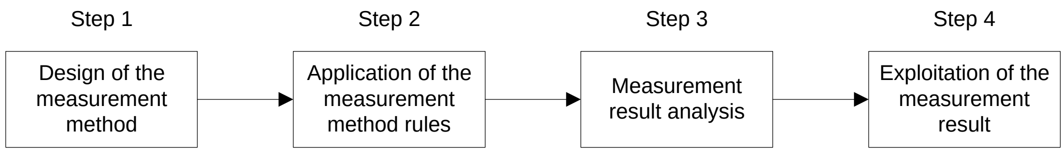


Figure 5: Measurement Process – High-level Model (Source: Abran & Jacquet, 1999)

It is to be noted that very few of the measurement concepts present in the ISO Vocabulary on Metrology address the first step (design of a measurement method) and none address the last step (exploitation of the measurements results) of the Abran and Jacquet process model [10]. This is illustrated in Table 3, which depicts a partial mapping between Figures 1 and 5: for instance, in [10,15], for the design of a measurement method, the Abran and Jacquet model includes more concepts than simply 'quantities and units'.

Table 3: Comparison of ISO [7] and the Abran & Jacquet Model [10]

Abran & Jacquet [10,15,17]	**Step 1** Design of Measurement Methods	**Step 2** Application of measurement method rules	**Step 3** Measurement results analysis	**Step 4** Exploitation of measurement results
ISO Categories of Metrology Terms [7]	• Quantities and units	• Measuring instruments • Characteristics of measuring instruments	• Measurement results	

4. Measurement steps and metrology concepts within SWEBOK

Using both the ISO set of metrology concepts model [11,12] and the measurement process model [10,15,17], we can analyze the current status of the field of 'software metrics' as documented in the Guide to the Software Engineering Body of Knowledge (SWEBOK) [13]. This SWEBOK project was initiated by the IEEE Computer Society to characterize the content of the Software Engineering Body of Knowledge and to consensually validate that portion of the Body of Knowledge that is both generally applicable and generally accepted. Over 450 experts from more than forty countries have participated to date in its review and validation, and the current version is currently being processed as an ISO software engineering technical report – ISO 19759 [16].
Table 4 presents first an inventory of the measurement-related statements appearing in two of the SWEBOK chapters [12]: software engineering management and software engineering quality.

Table 4: Inventory of measurement-related statements in two SWEBOK chapters

SWEBOK chapters and sections	Measurement-related statements
Chapter: Software Engineering Management	
Software Engineering Measurement	
Goals (p. 8-7)	Determining the goals of a measurement program *Ad hoc* approach to software engineering measurement characterized early efforts Organizational objectives Software process improvement goals
Measurement Selection (p. 8-8)	Goal-driven measurement selection Measurement validity
Measuring Software and its Development (p. 8-8)	Size measurement Structure measurement Resource measurement Quality measurement
Collection of data (p.8-9)	Survey techniques and form design Automated and manual data collection
Software Measurement Models (p. 8-9)	Model building, calibration and evaluation Implementation, interpretation and refinement of models
Chapter: Software Engineering Quality	
Software Quality Concepts	
Measuring the value of quality (p. 11-2)	Determination of a value of a software project
Measuring Applied to Software Quality Assurance SQA and Verification & Validation V&V	
Fundamentals of Measurement (p. 11-10)	Theory of measurement Measurement scales Measurement programs are useful if they help project stakeholders: Understand what is happening during their processes Control what is happening on their projects Measurement practices: experimentation and data collection
Measures (p. 11-11)	Measurement models and framework for software quality Types of measures
Measurement analysis techniques (p.11-11)	Mathematical and graphical techniques Statistics-based techniques and test
Defect characterization (p.11-11)	Defect taxonomies Analyzing defects Measurement approaches
Additional Uses of SQA and V&V data (p. 11-12)	Determine how the SQA and V&V processes use measurement directly to support achieving their goals Reliability models and benchmarks

Table 5 lists, for each of the ten chapters of SWEBOK, which metrology concepts and measurement steps are addressed whenever a measurement-related statement appears in this Guide. It can be observed that a large majority of the measurement-related concepts mentioned in SWEBOK are listed in the category of concepts related to the exploitation of the measurement results. Very few SWEBOK statements directly address the measuring instrument or the quality of the direct measurement results (prior to their use in quantitative analytical models (assessment models or predictive models)), and only one in the Software Quality chapter addresses a single aspect of the design of measurement instrument, through a subset of the concepts of quantities and units. Further work is in progress aimed at a more in-depth study of each measurement-related statement in all SWEBOK chapters, which also includes an analysis of the seminal references quoted in each chapter dealing with measurement-related concepts.

Table 5: Measurement steps and metrology category of concepts within SWEBOK [13]

Abran et Jacquet [10,15,17]	**Step 1** Design of Measurement Methods	**Step 2** Application of measurement method rules	**Step 3** Measurement results analysis	**Step 4** Exploitation of measurement results
ISO Metrology Vocabulary [7]	- Quantities and units	-Measuring instruments -Characteristics of measuring instruments	- Measurement results	
Software Engineering Requirements				
Process quality and improvement				×
Requirements negotiation				×
Document quality				×
Acceptance tests				×
Requirements tracing				×
Software Engineering Design				
Measures			×	
Software Engineering Testing				
Evaluation of the program under test				×
Evaluation of the tests performed				×
Software Engineering Maintenance				
Software Maintenance Measurement				×
Software Configuration Management (SCM)				
Surveillance of software configuration management				×
Software Engineering Management				
Goals				×
Measurement Selection				×
Measuring Software and its Development				×
Collection of data		×		
Software Measurement Models			×	
Software Engineering Process				
Methodology in process measurement		×		
Process Measurement Paradigms				×
Software Engineering Quality				
Measuring the value of quality				×
Fundamentals of Measurement	×			
Measures			×	
Measurement analysis techniques				×
Defect characterization				×
Additional Uses of SQA and V&V data				×

5. Observations

While 'software metrics' are most often proposed as the measurement tools of choice in empirical studies in software engineering, this field of 'software metrics' has most often been discussed from the perspective referred to as ‘measurement theory’. However, in other disciplines, it is the domain of knowledge referred to as ‘metrology’ that is the foundation for the development and use of measurement instruments and measurement processes. In this position paper, we have used our initial modeling [11,12] of the sets of measurement concepts documented in the ISO International Vocabulary of Basic and General Terms in Metrology to survey, and position, the measurement-related statements in the Guide to the Software Engineering Body of Knowledge. This has revealed that, even though measurement-related statements appear throughout the SWEBOK document, they overwhelmingly concern the use of measurement results in assessment and predictive models. By contrast, there is very little in the document relating to the quality of the quantitative inputs to these models, and almost nothing about the supporting measuring instruments necessary to obtain these inputs. This illustrates that, in the software engineering literature, there is as yet very little discussion, or related consensus, on the topic of measuring instruments so overwhelmingly present in the traditional engineering disciplines and culture. This also illustrates that most of the metrology concepts, and sub-concepts have not yet been extensively discussed or addressed in the 'software metrics' literature. In the context where measuring instruments are necessary key elements of empirical studies, this points to a potentially significant weakness in current empirical studies in software engineering, while at the same time providing an indication of where metrology-related improvements in software measurement could contribute significantly to strengthening future empirical studies in software engineering.

6. References

[1] A. Abran, R. Dumke, ‘A Framework of the Software Measurement Domain: Current Situation and Future Directions’, International Workshop on Software Measurement – IWSM 97, 1997.

[2] S. R. Chidamber, C. F. Kemerer, ‘A Metrics Suite for Object Oriented Design’, IEEE Transactions on Software Engineering. Vol. 20, No. 6, June 1994, pp 476-493.

[3] E. Dimitrov, R, R. Dumke, E. Foltin, M. Wipprecht, ‘Conception and Experience of Metrics Based Software Reuse in Practice’, International Workshop on Software Measurement (IWSM) Lac Supérieur, Québec, 1999.

[4] T. Fetcke, A. Abran, R, A. Dumke, ‘Generalized Representation for Functional Size Measurement Methods’, International Workshop on Software Measurement – ISWM 2001, Montréal, Canada, August 28-29, 2001.

[5] R. Dumke, E. Foltin, P. Bourque, ‘Issues when Evaluating Software Measurement Tools Based on Existing Standards’, Forum on Software Engineering Standards (SES’96), Montréal, IEEE-CS Press, 1996.

[6] R. Dumke, M. Lother, A. Abran, ‘An Approach for Integrated Software Measurement Processes in the IT Area’, 4th European Conference on Software Measurement and ICT Control in cooperation with DASMA (FESMA-DASMA), Heidelberg, Germany, 2001.

[7] ISO, ISO: International Vocabulary of Basic and General Terms in Metrology, International Organization for Standardization -ISO, Geneva, 1993.

[8] Zuse, H., A Framework of Software Measurement, New York, 1997.

[9] M. Hitz, B. Montazeri, ‘Chidamber and Kemerer's Metrics Suite: A Measurement Theory Perspective’, IEEE Transactions on Software Engineering. Vol. 22, No. 4, pp 267-271, April 1996.

[10] Abran, A., Jacquet, J.-P., A Structured Analysis of the New ISO Standard on: "Functional Size Measurement - Definition of Concepts" (ISO/IEC 14143-1), 4th IEEE International Software Engineering Standards Symposium, ISESS'99, Curitiba, Brazil, IEEE-CS Press,1999.

[11] A. Abran, Sellami. A., ‘Initial Modelling of the Measurement Concepts in the ISO Vocabulary of Terms in Metrology’, 12th International Workshop on Software Measurement - ISWM 2002, Magdeburg, Germany, 2002.

[12] A. Sellami, Traitement de la mesure dans le guide SWEBOK, lecture dirigée, École de technologie supérieure, Montréal, Canada, August 2002.

[13] Abran, A., Moore, J., Bourque, P., Dupuis, R. L. Tripp, L., Guide to the Software Engineering Body of Knowledge – SWEBOK, Trial Version, IEEE-Computer Society Press, Dec. 2001, URL: http://www.swebok.org

[14] ISO *ISO/IEC FDTR 14143-3* – Information technology – Software measurement – Functional size measurement – Verification Guide, International Organization for Standardization – ISO, Geneva, to be published, 2002.

[15] Jacquet, J.P, Abran, A. ‘From Software Metrics to Software Measurement Methods: A Process Model’, Third International Symposium and Forum on Software Engineering Standards (ISESS'97), Walnut Creek, CA, IEEE-CS Press, 1997.

[16] ISO, FDTR: 19759 Guide to the Software Engineering Body of Knowledge – SWEBOK, International Organization for Standardization – ISO, Geneva, to be published 2002.

[17] Jacquet, J.-P., Abran, A., *Metrics Validation Proposals: A Structured Analysis*, presented at 8th International Workshop on Software Measurement, Magdeburg, Germany, 1998.

Appendix: Sample of figures on metrology-related concepts as modeled in [11,12]

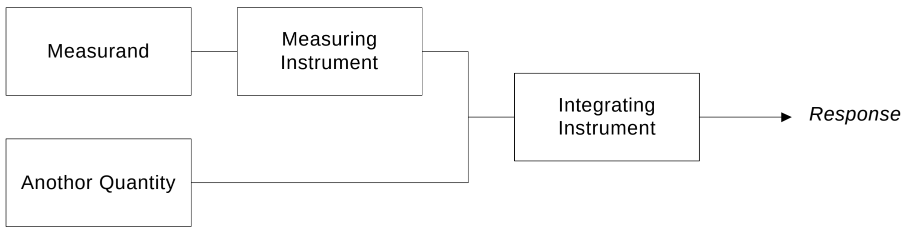


Figure A.1: Integrating Instrument

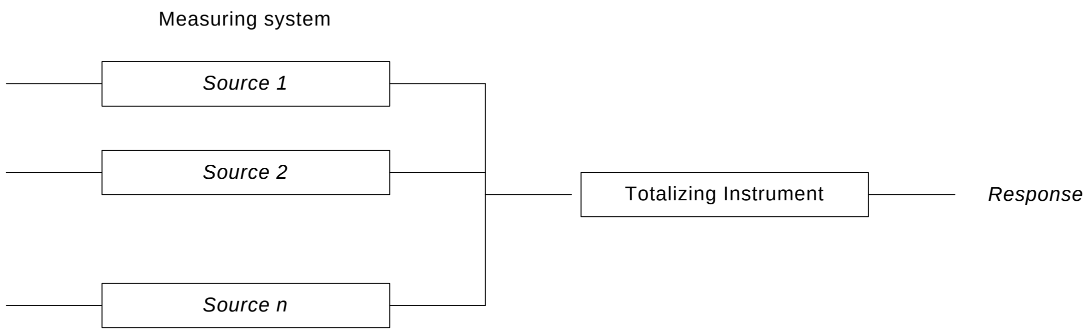


Figure A.2: Model of a totalizing measuring instrument

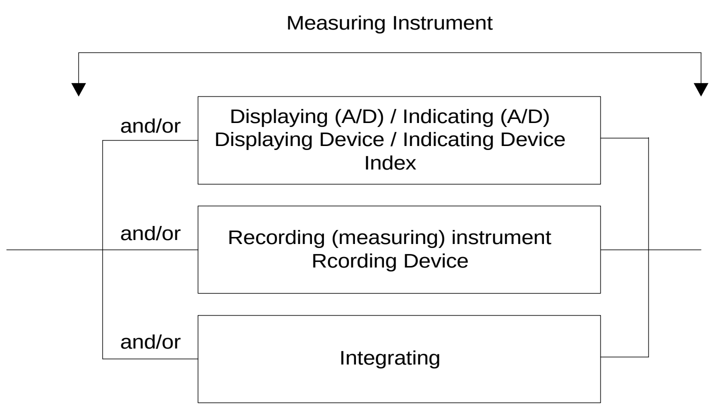


Figure A.3: Details of a 'Measuring Instrument'

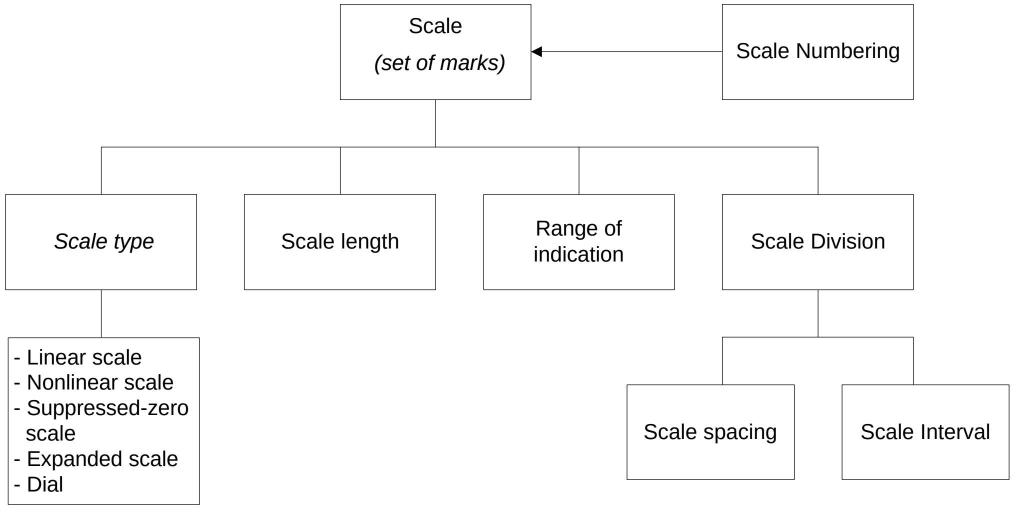


Figure A.4: Detailed topology of 'Scale'