Abran, Alain.; Al-Qutaish, Rafa E. and Cuadrado-Gallego, Juan, "Investigation of the Metrology Concepts in ISO 9126 on Software Product Quality Evaluation", In *Proceedings of the 10th WSEAS International Conference on Computers (ICComp'2006)*, July 13-15, Athens, Greece, 2006, pp. 864-872. *(ISBN: 960-8457-47-5)*

Investigation of the Metrology Concepts in ISO 9126 on Software Product Quality Evaluation

ALAIN ABRAN*, RAFA E. AL-QUTAISH*, JUAN J. CUADRADO-GALLEGO**
*École de Technologie Supérieure
University of Québec
1100 Notre-Dame Street West
Montréal, Québec H3C 1K3,
CANADA
** Department of Computer Science
University of Alcalá
28805 Edificio Politécnico
Alcalá de Henares,
SPAIN
alain.abran@etsmtl.ca, rafa.al-qutaish.1@ens.etsmtl.ca, jjcg@uah.es

Abstract: - The ISO International Vocabulary of Basic and General Terms in Metrology (VIM) represents the international consensus on a common and general terminology of metrology concepts. However, until recently, it was not usual practice in software engineering measurement to take into account metrology concepts and criteria in the design of software measures. Using the ISO 9126-4 Technical Report on the measurement of software quality in use as a case study, this paper reports on the extent to which this ISO series addresses the metrology criteria typical of classic measurement. Areas for improvement in the design and documentation of measures proposed in ISO 9126 are identified.

Key-Words: - Metrology, VIM, ISO 9126, Quality in Use metrics, Software Measurement, Software Quality.

1 Introduction

In the field of software engineering, the term "metrics" is used in reference to multiple concepts, whether in terms of the quantity to be measured (measurand[1]), measurement procedures, measurement results or models of relationships across multiple measures, or of the objects themselves. In the software engineering literature, the term is applied, for instance, to a measure of a concept (e.g. McCabe cyclomatic complexity [1]), to quality models (ISO 9126 – software product quality [2]) and to estimation models (e.g. Halstead's equations [3], COCOMO I and II estimation models [4, 5]). This has led to many curious problems, among them a proliferation of numerous publications on metrics for concepts of interest, but with a very low rate of acceptance and use by either researchers or practitioners, as well as a lack of consensus on how to validate so many proposals [6-8]. The inventory of software metrics is at the present time so diversified and includes so many individual proposals [9, 10] that it is not seen to be economically feasible for either industry or the research community to investigate each of the hundreds of alternatives proposed to date.

While metrology has a long tradition of use in physics and chemistry, it is rarely referred to in the software engineering literature. Carnahan et al. [11] are among the first authors to identify this gap in what they referred to as "IT metrology"; they highlight the challenges and opportunities arising from the application of the metrology concepts to information technology. In addition, they have proposed logical relationships between metrology concepts, consisting of four steps to follow to obtain measured values: defining quantity/attribute, identifying units and scales, determining the primary references and settling the secondary references. Moreover, Gray [12] discusses the applicability of metrology to information technology from the software measurement point of view.

Abran [13] has highlighted some high-level ambiguities in the domain of software measurement,

[1] A measurand is defined as a particular quantity subject to measurement; the specification of a measurand may require statements about quantities such as time, temperature and pressure [19].

and proposed substituting the appropriate metrology terms for the current ambiguous and peculiar software metrics terminology unique to the domain of software engineering. In metrology, the term "metrics" is never used. In addition, the availability of the metrology concepts in software engineering has been investigated in [7, 14, 15]. Abran and Sellami [8] have documented the metrology concepts addressed in ISO 19761 (COSMIC-FFP), both in the design of this measurement method and in some of its practical uses. Moreover, Sellami and Abran [16] have investigated the contribution of metrology concepts to understanding and clarifying the framework for software measurement validation proposed by Kitchenham et al. in [17].

The ISO 9126 series of documents on software product quality evaluation proposes a set of 120 metrics[2] for measuring the various characteristics and subcharacteristics of software quality. However, as is typical in the software engineering literature, their set of so-called metrics in ISO 9126 refers to multiple distinct concepts which, in metrology, would have distinct labels (or naming conventions, e.g. terms) to avoid ambiguities.

To help in understanding and clarifying the nature of the metrics proposed in ISO TR 9126-4 [18], each is analyzed in this paper from a metrology perspective and mapped to the relevant metrology concepts. Such a mapping will also contribute to identifying the measurement concepts that have not yet been tackled in the ISO 9126 series of documents. Each of these gaps represents an opportunity for improvement in the design and documentation of the measures proposed in ISO 9126.

This paper presents an overview of the relevant metrology concepts in section 2, and an overview of the ISO 9126 series and the quality in use metrics in Section 3. Section 4, 5, 6, and 7 present the analysis of the "effectiveness", "productivity", "safety", and "satisfaction" metrics, respectively. A discussion in section 8 concludes the paper.

2 Metrology

The term "metrology" is defined in the ISO International Vocabulary of Basic and General Terms in Metrology as the field of knowledge dealing with measurement [19]. More specifically, it has been defined in [20] as: "*that portion of measurement science used to provide, maintain, and disseminate a consistent set of units; to provide support for the enforcement of equity in trade by weights and measurement laws; or to provide data for quality control in manufacturing*". Following the above definitions, metrology forms the basis of all measurement-related concepts in the natural sciences and engineering, and to each of the different interpretations of a software metrics is associated a related distinct metrology term with related metrology criteria and relationships with other measurement concepts. In 1984, the ISO, with other participating organizations (BIPM, IEC and OIML), published their first edition of the international consensus on the basic and general terms in metrology (VIM) [21]. Later, in 1993, this publication was reviewed, and then the ISO, in collaboration with six participating organizations (BIPM, IEC, OIML, IUPAC, IUPAP and IFCC), published the second edition of this document [19]. The ISO is now working on its third edition of this document to integrate, in particular, concepts related to measurement uncertainty and measurement traceability.

The second VIM edition on metrology presents 120 terms in six categories and in increasing order of complexity, and describes each term individually in textual format (in parentheses, the number of terms in each category): Quantities and Units (22 terms), Measurements (9 terms), Measurement Results (16 terms), Measuring Instruments (31 terms), Characteristics of the Measuring Instruments (28 terms) and Measurement Standards – Etalon (14 terms).

To facilitate an understanding of these more than one hundred related terms, Abran and Sellami [22] proposed a modeling of all the sets of measurement concepts documented in this ISO document.

Two of the categories of terms in the VIM deal with some aspects of the design of measurement methods, that is, category 1: "quantities and units", and category 2: "measurement standards – etalon". The other four categories are related to the application of a measurement design with a measuring instrument, and to the quality characteristics of the measurement results provided by this measuring instrument [22]. More specifically, we use the first category, which deals with the design of the measurement methods, that is, quantities and units and, in particular, the system of

[2] While the term "metrics" is used in ISO 9126, the use of this term will be abandoned and replaced by "measures" in the next ISO version currently in preparation as an initial step towards harmonizing the software engineering measurement terminology with the metrology terminology.

quantities that consists of two types of quantities; that is, base and derived.[3]

3 ISO 9126 & Quality in Use Metrics

In 1991, the ISO published its first international consensus on the terminology for the quality characteristics for software product evaluation: ISO 9126 – *Software Product Evaluation – Quality Characteristics and Guidelines for their Use* [23]. From 2001 to 2004, the ISO published an expanded four-part version, containing both the ISO quality models and inventories of proposed measures for these models. The current version of the ISO 9126 series now consists of one International Standard (IS) [2] and three Technical Reports (TR) [18, 24, 25].

The first document of the ISO 9126 series – *Software Product Quality Model* – contains two structures of quality models for software product quality [2]: a structure for both the internal and external quality models, and a structure for the quality in use model. The first structure of the ISO 9126-1 Quality Model includes six characteristics, subdivided into twenty-seven subcharacteristics for internal and external quality [2]. These subcharacteristics are related to internal software attributes, and are noticeable externally when the software is used as part of a computer system. The second structure of the ISO 9126-1 model includes four quality in use characteristics [2]: effectiveness, productivity, safety and satisfaction. These characteristics and subcharacteristics are defined in the ISO 9126-1 international standard document.

The second document of the ISO 9126 series – *Software Product External Quality Metrics* – contains a set of metrics for each external quality subcharacteristic, explanations of how to apply and use them, and examples of how to apply them during the software product life cycle [24].

The third document of the ISO 9126 series – *Software Product Internal Quality Metrics* – contains an inventory of metrics for each internal quality subcharacteristic, explanations of the application of these metrics, and examples of how to use them in the software product life cycle [25].

Finally, the fourth document of the ISO 9126 series – *Software Product "Quality in Use" Metrics* – contains a basic set of metrics for each quality in use characteristic, explanations of how to apply them and examples of how to use them in the software product life cycle [18].

In ISO 9126-4 [18], fifteen metrics have been proposed for the quality in use metrics. They have been classified into four collections of metrics based on the characteristics presented in ISO 9126-1:

1. Effectiveness: task effectiveness, task completion and error frequency
2. Satisfaction: task time, task efficiency, economic productivity, productive proportion and relative user efficiency
3. Safety: user health and safety, safety of people affected by use of the system, economic damage and software damage
4. Productivity: satisfaction scale, satisfaction questionnaire and discretionary usage

These fifteen metrics are analyzed using a metrology concept structure from the VIM category, *Quantities and Units* [19], based on four characteristics, that is: system of quantities, dimension of a quantity, unit of measurement and value of a quantity.

4 Effectiveness Metrics

In ISO 9126-4, the claim is that the three *Effectiveness Metrics* assess whether or not the task carried out by users achieved the specific goals with accuracy and completeness in a specific context of use [18]. This section presents the outcomes of the mapping of the set of *Quantities and Units* metrology concepts to the 2004 description of *Effectiveness Metrics* in ISO 9126-4. A summary of this mapping is presented in the Appendix.

4.1 System of quantities for Effectiveness

4.1.1 Base quantities

First, it can be observed that these three *Effectiveness Metrics* are not collected directly by a measurement system, but are derived from a computation of four base quantities that are themselves collected directly, that is: task time, number of tasks, number of errors made by the user and proportional value of each missing or incorrect component.

The first three of these base measures in the Appendix refer to terms in common use, but this leaves much to interpretation on what constitutes, for example, a task: it does not ensure that the measurement results are repeatable and reproducible across measurers, across groups measuring the same software and, as well, across organizations where a task might be interpreted differently and with

[3] In ISO 15939, the terms "base quantities" and "derived quantities" were replaced by equivalent terms: "base measures" and "derived measures".

different levels of granularity. This leeway in their interpretation makes a rather weak basis for either internal or external benchmarking.

The third base quantity, *number of errors made by the user*, is defined in Appendix F of ISO TR 9126-4 as an "instance where test participants did not complete the task successfully, or had to attempt portions of the task more than once" [18]. This definition diverges significantly from the one in the IEEE Standard Glossary of Software Engineering Terminology [26] where the term "error" has been defined as "the difference between a computed, observed, or measured value or condition and the true, specified, or theoretically correct value or condition; for example, a difference of 30 meters between a computed result and the correct result."

The fourth base quantity, referred to as the "Proportional value of each missing or incorrect component" in the task output is based, in turn, on another definition, whereas each "potential missing or incorrect component" is given a weighted value A_i based on the extent to which it detracts from the value of the output to the business or user [18]. These expansive, embedded definitions contain a number of subjective assessments for which no repeatable procedure is provided: the value of the output to the business or user, the extent to which it detracts, the components of a task and potential missing or incorrect components.

4.1.1 Derived quantities

The proposed three *Effectiveness Metrics*, which are defined as a prescribed combination of these base quantities, are therefore derived quantities. The ranges of the results obtained from implementing the corresponding measurement function are introduced in the upper part of the Appendix for each of these derived quantities. These quantities inherit the weaknesses of the base quantities of which they are composed.

4.2 Dimension of a quantity for Effectiveness

Emerson [27] states that the concept of dimension is particularly applicable to the derived quantities: two of them, i.e. task effectiveness and task completion, can have values between 0 and 1, and would be considered as dimensionless quantities, since a ratio of quantities with the same dimensions is itself dimensionless [27].

4.3 Units of measurement for Effectiveness

The metrology concepts related to units of measurement are:

- Symbols of the units
- Systems of units
- Coherent (derived) units
- Coherent system of units
- International system of units
- Base units
- Derived units
- Off-system units
- Multiples of a unit
- Submultiples of a unit

The mappings of these metrology concepts for *Effectiveness Metrics* are presented in the Appendix. Two metrology concepts must be analyzed in more detail, *base units* and *derived units*.

4.3.1 Base units

Of the four base quantities, a single one, i.e. task time, has an internationally recognized standard base unit, i.e. the second, or a multiple of this unit. It also has a universally recognized corresponding symbol ('s'). The next two base units (tasks and errors) do not refer to any international standard of measurement, and must be locally defined (which means that they fit poorly, for comparison purposes, when measured by different people, unless local measurement protocols have been clearly documented, and they are implemented rigorously in a specific organization). The fourth base quantity, proportional value of each missing or incorrect component, is puzzling because it is based on a given weighted value (number) and has no measurement unit.

4.3.2 Derived units

The derived quantity, task effectiveness, leads to a derived unit that depends on a given weight (i.e. (1 – a given weight)). Therefore, like the base unit, its derived unit of measurement is unclear.

The derived quantity, task completion, is computed by dividing two base quantities (task/task) with the same unit of measurement.

The definition of the computation of the derived quantity, error frequency, provides two distinct alternatives for the elements of this computation. This can lead to two distinct interpretations, i.e. errors/task or errors/second. Of course, this in turn leads to two distinct derived quantities as a result of implementing two different measurement functions (formulas) for this derived quantity. Of course, this leaves open the possibility of misinterpretation and misuse of measurement results when combined with other units: for example, measures in centimeters and measures in inches cannot be added or multiplied.

This lack of clarity, as well as the lack of references to international units of measurement, could explain why there has been no attempt to integrate the proposed base and derived quantities into a system of units, including references to coherent units and a coherent system of units.

4.4 Value of a quantity for Effectiveness

The four types of metrology values of a quantity are: true value, conventional true value, numerical value and conventional reference scale.

Numerical values are indeed obtained for each base quantity based on the defined data collection procedure; for each derived quantity, the numerical values are obtained by applying their respective measurement function. For instance, the derived quantities, task effectiveness and task completion, are both percentages, and are interpreted as the effectiveness and completion of a specific task respectively.

For task effectiveness in particular, anyone would be hard pressed to figure out both a true value and a conventional true value; for task completion and error frequency, the true values would depend on locally defined and rigorously applied measurement procedures, but without reference to universally recognized conventional true values (as they are locally defined).

Finally, in terms of the metrological values of a quantity, only task time refers to a conventional reference scale, that is, the international standard-etalon for time, from which the second is derived. None of the other base quantities in these effectiveness metrics refers to a conventional reference scale, or to a locally defined one.

5 Productivity Metrics

In ISO 9126-4, the claim is made that the five productivity metrics assess the resources that users consume in relation to the effectiveness achieved in a specific context of use. In this standard, the time required to complete a task is considered to be the main resource to take into account [18]. This section presents the outcome of the mapping of the set of *Quantities and Units* metrology concepts to the 2001 description of *Productivity Metrics* in ISO 9126-4.

5.1 System of quantities for Productivity

One of the five proposed productivity metrics in ISO 9126-4 is a base quantity (task time) while the other four are derived quantities (task efficiency, economic productivity, productive portion and relative user efficiency).

In addition, task efficiency refers explicitly to another derived quantity, task effectiveness, which was analyzed in the previous section.

It is to be noted that these derived quantities are themselves based on five base quantities: task time, cost of the task, help time, error time and search time.

5.2 Dimension of a quantity for Productivity

All the productivity metrics, except task time, are dimensionless quantities.

5.3 Units of measurement for Productivity

For the base and derived quantities, there are five base units and no explicit derived units. However, it can be observed that the measurement unit for task effectiveness is not completely clear, since it depends on an ill-defined “given weight”:

$$\text{'task efficiency' unit} = \frac{\text{'task effectiveness' unit}}{\text{second}}.$$
$$= \frac{\text{1 - 'a given weight' unit}}{\text{second}} = \frac{?}{\text{second}}. \quad (1)$$

Similarly, the measurement unit of economic productivity depends on the measurement unit of task effectiveness, a derived quantity which is unknown:

$$\text{'economic productivity' unit} = \frac{\text{'task effectiveness' unit}}{\text{currency unit}}$$
$$= \frac{\text{1 - 'a given weight' unit}}{\text{currency unit}}$$
$$= \frac{?}{\text{currency unit}}. \quad (2)$$

Since there is no measurement unit for the productive proportion (it has the same measurement unit in both the numerator and the denominator), the result is a percentage:

$$\text{'productive proportion' unit} = \frac{\text{second}}{\text{second}}. \quad (3)$$

Finally, for relative user efficiency, there is no measurement unit either, since the measurement units in both the numerator and the denominator are the same here as well (the task efficiency measurement unit), and therefore the result of this derived quantity is also a percentage. This point can be clarified as follows:

$$
\text{'relative user efficiency' unit} = \frac{\text{'task efficiency' unit}}{\text{'task efficiency' unit}} = \frac{\frac{\text{'task effectiveness' unit}}{\text{second}}}{\frac{\text{'task effectiveness' unit}}{\text{second}}} = \frac{\frac{1-\text{'a given weight' unit}}{\text{second}}}{\frac{1-\text{'a given weight' unit}}{\text{second}}} = \frac{\frac{?}{\text{second}}}{\frac{?}{\text{second}}}. \quad (4)
$$

6 Safety Metrics

In ISO 9126-4, the safety metrics claim to assess the level of risk of harm to people, businesses, software, property or the environment in a specific context of use; their scope includes the health and safety of both the users and those who affected by use, as well as unintended physical or economic consequences [18].

To evaluate the safety characteristics of a software product, four derived quantities must be quantified (i.e. user health and safety, software damage, economic damage and the safety of people affected by use of the system). Each of these derived quantities is the result of a computational formula (function), which consists of a combination of pre-collected base quantities (i.e. number of usage situations, number of people, number of occurrences of software corruption, number of occurrences of economic corruption and number of users). It can be observed that the resulting values of all the derived quantities should be between 0 and 1.

All the safety metrics are dimensionless quantities; there are five base units and two derived units for these quantities. In addition, two of the derived quantities have no measurement units, since the measurement unit is the same in both the numerator and the denominator, i.e. user health and safety and safety of people affected by use of the system, whereas none of the measurement units has a symbol.

7 Satisfaction Metrics

The satisfaction metrics in ISO 9126-4 claim to assess the user's attitudes towards the use of the product in a specific context of use [18].

All three proposed satisfaction metrics are derived quantities (i.e. satisfaction scale, satisfaction questionnaire and discretionary usage), which themselves depend on four base quantities (i.e. population average, number of responses, number of times that specific software function / application / systems are used and number of times that specific software function/application/systems are intended to be used). Two of the proposed satisfaction metrics are dimensionless quantities, i.e. satisfaction questionnaire and discretionary usage.

Regarding the measurement units, there are four base units and no derived units; however, the measurement unit, satisfaction scale, is not clear, since it depends on a "questionnaire producing psychometric scales". The clarification of this point is as follows:

$$
\text{'satisfaction scale' unit} = \frac{\text{psychometric scale unit}}{\text{people}}. \quad (5)
$$

8 Conclusions

The ISO International Vocabulary of Basic and General Terms in Metrology (VIM) represents the international consensus on a common and general terminology of metrology concepts. However, until recently, it was not usual practice in software engineering measurement to take into account metrology concepts and criteria, either in the design of software measures or in their use and in the interpretation of measurement results.

This paper has presented an analysis of the ISO 9126-4 Technical Report on quality in use metrics, and has investigated the extent to which it addresses the metrology criteria found in classic measurement. Based on the analysis in sections 4 to 7, the following comments and suggestions can be made:

- Identifying and classifying the quality in use metrics into base and derived quantities makes it easy to determine which should be collected (base quantities) to be used subsequently in computing the other (derived) quantities.
- Based on equations (1) and (3 to 5), some of the derived units are ambiguous, since they depend on other quantities with unknown units.
- None of the quality in use metrics refers to any system of units, coherent (derived) unit, coherent system of units, international system of units (SI), off-system units, multiple of a unit, submultiple of a unit, true values, conventional true values or numerical values.
- None of the base and derived quantities, except for task time, has symbols for their measurement units.

It is to be noted that the ranges of the results of many of the derived metrics in ISO 9126-4 are between 1 and 0. Therefore, it is easy to convert them to percentage values. However, from our point of view, these results will be easier to understand if they are ranked in terms of qualitative values; for example, for task completion, if the percentage result is 100%, then the completion of the task is labeled “excellent”; if the result is 80%, then the completion of the task is labeled “very good”; and so on.

Using the ISO 9126-4 Technical Report on the measurement of software quality in use as a case study, this paper has investigated and reported on the extent to which this ISO series addresses the metrology criteria typical of classic measurement. Areas for improvement in the design and documentation of measures proposed in ISO 9126 have been identified. The analysis methodology developed to investigate ISO TR9126-4 could also be of use to analyze the metrological strengths and weaknesses of close to 120 metrics proposed by the ISO in TRs 9126-2 and -3.

References:

[1] McCabe, T. J., "A Complexity Measure," *IEEE Transaction on Software Engineering*, Vol. 2, No. 4, 1976, pp. 308-320.

[2] ISO/IEC, *ISO/IEC 9126-1: Software Engineering - Product Quality - Part 1: Quality Model*, International Organization for Standardization, Geneva, Switzerland, 2001.

[3] Halstead, M. H., *Elements of Software Science*, Elsevier North-Holland, New York, 1977.

[4] Boehm, B., *Software Engineering Economics*, Prentice-Hall, Upper Saddle River, NJ, USA, 1981.

[5] Boehm, B., Clark, B., Horowitz, E., Westland, C., Madachy, R., and Selby, R., "Cost models for future software life cycle processes: COCOMO 2.0," *Annals of Software Engineering*, Vol. 1, No. 1, 1995, pp. 57-94.

[6] Jacquet, J. P. and Abran, A., "Metrics Validation Proposals: A structured Analysis," in Proceedings of the 8th International Workshop on Software Measurement, Magdeburg, Germany, 1998.

[7] Abran, A. and Sellami, A., "Measurement and Metrology Requirements for Empirical Studies in Software Engineering," in Proceedings of the 10th International Workshop on Software Technology and Engineering Practice (STEP), Montreal, Canada, 2002.

[8] Abran, A. and Sellami, A., "Analysis of Software Measures Using Metrology Concepts - ISO 19761 Case Study," presented in the International Workshop on Software Audits and Metrics SAM'2004, 6th International Conference on Enterprise Information Systems ICEIS 2004, Porto, Portugal, 2004.

[9] Kaner, C. and Bond, W. P., "Software Engineering Metrics: What Do They Measure and How Do We know?" in Proceedings of the 10th International Software Metrics Symposium (METRICS 2004), Chicago, Illinois, USA, 2004.

[10] Zuse, H., *A Framework of Software Measurement*, Walter de Gruyter, Berlin, Germany, 1998.

[11] Carnahan, L., Carver, G., Gray, M., Hogan, M., Hopp, T., Horlick, J., Lyon, G., and Messina, E., "Metrology for Information Technology," *StandardView*, Vol. 5, No. 3, 1997, pp. 103-109.

[12] Gray, M. M., "Applicability of Metrology to Information Technology," *Journal of Research of the National Institute of Standards and Technology*, Vol. 104, No. 6, 1999, pp. 567-578.

[13] Abran, A., "Software Metrics Need to Mature into Software Metrology (Recommendations)," presented in the NIST Workshop on Advancing Measurements and Testing for Information Technology (IT), Maryland, USA, 1998.

[14] Abran, A., Sellami, A., and Suryn, W., "Metrology, Measurement and Metrics in Software Engineering," in Proceedings of the 9th International Software Metrics Symposium, Sydney, Australia, 2003.

[15] Bourque, P., Wolff, S., Dupuis, R., Sellami, A., and Abran, A., "Lack of Consensus on Measurement in Software Engineering: Investigation of Related Issues," in Proceedings of the 14th International Workshop on Software Measurement (IWSM), Magdeburg, Germany, 2004, pp. 321-333.

[16] Sellami, A. and Abran, A., "The contribution of metrology concepts to understanding and clarifying a proposed framework for software measurement validation," in Proceedings of the 13th International Workshop on Software Measurement (IWSM), Montreal, Canada, 2003, pp. 18-40.

[17] Kitchenham, B., Pfleeger, S. L., and Fenton, N., "Towards a Framework for Software Measurement Validation," *IEEE Transaction on Software Engineering*, Vol. 21, No. 12, 1995, pp. 929-944.

[18] ISO/IEC, *ISO/IEC TR 9126-4: Software Engineering - Product Quality - Part 4: Quality in Use Metrics*, International Organization for Standardization, Geneva, Switzerland, 2004.

[19] ISO/IEC, *International Vocabulary of Basic and General Terms in Metrology (VIM)*, International Organization for Standardization, Geneva, Switzerland, 1993.

[20] Simpson, J. A., "Foundations of Metrology," *Journal of Research of the National Bureau of Standards*, Vol. 86, No. 3, 1981, pp. 36-42.

[21] ISO/IEC, *International Vocabulary of Basic and General Terms in Metrology (VIM)*, International Organization for Standardization, Geneva, Switzerland, 1984.

[22] Abran, A. and Sellami, A., "Initial Modeling of the Measurement Concepts in the ISO Vocabulary of Terms in Metrology," in Proceedings of the 12th International Workshop on Software Measurement (IWSM), Magdeburg, Germany, 2002.

[23] ISO/IEC, *ISO/IEC IS 9126: Software Product Evaluation - Quality Characteristics and Guidelines for Their Use*, International Organization for Standardization, Geneva, Switzerland, 1991.

[24] ISO/IEC, *ISO/IEC TR 9126-2: Software Engineering - Product Quality - Part 2: External Metrics*, International Organization for Standardization, Geneva, Switzerland, 2003.

[25] ISO/IEC, *ISO/IEC TR 9126-3: Software Engineering - Product Quality - Part 3: Internal Metrics*, International Organization for Standardization, Geneva, Switzerland, 2003.

[26] IEEE, *Std. 610.12-1990: Standard Glossary of Software Engineering Terminology*, the Institute of Electrical and Electronics Engineers, New York, USA, 1990.

[27] Emerson, W. H., "Short Communication on the Concept of Dimension," *Metrologia*, Vol. 42, No. 2, 2005, pp. 21–22.

Appendix: “Quantities and Units” Metrology Concepts in ISO 9126-4 – Effectiveness Metrics

Metrology Concepts	ISO 9126-4 (Effectiveness Metrics*)
5.1 System of Quantities:	
- Base Quantities:	1. Task Time. 2. Number of Tasks. 3. Number of Errors Made by the User. 4. Proportional value of each missing or incorrect component.
- Derived Quantities:	5. Task Effectiveness*. $0 \leq$ Task Effectiveness ≤ 1 6. Task Completion*. $0 \leq$ Task Completion ≤ 1 7. Error Frequency*. Error Frequency ≥ 0
5.2 Dimension of a Quantity:	
- Quantities of Dimension One (Dimensionless Quantities):	5. Task Effectiveness. 6. Task Completion.
5.3 Units of Measurement:	
- Symbols of the Units:	- *s* (Second)
- Systems of Units:	- None.
- Coherent (Derived) Units:	- None.
- Coherent System of Units:	- None.
- International System of Units (SI):	- None.
- Base Units:	1. Second. 2. Task. 3. Error. 4. Non (*ill-defined*)
- Derived Units:	5. (1- a given weight). 6. Task/Task = % 7. Error/Task or Error/Second.
- Off-System Units:	- None.
- Multiple of a Unit:	- None.
- Submultiple of a Unit:	- None.
5.4 Value of a Quantity:	
- True Values:	- None.
- Conventional True Values:	- None.
- Numerical Values:	- Results of applying the measurement functions of the above base and derived quantities.
- Conventional Reference Scales (Reference-Value Scales):	1- Task Time.