

# Attributes and Related Base Measures Within ISO 9126: A Pareto Analysis

Jean-Marc Desharnais   Alain Abran   Witold Suryn

École de technologie supérieure – Université du Québec  
1100 Notre-Dame, Montréal, Québec, Canada, H3C 1K3

[desharnaisjm@gmail.com](mailto:desharnaisjm@gmail.com)   [alain.abran@etsmtl.ca](mailto:alain.abran@etsmtl.ca)   [witold.suryn@etsmtl.ca](mailto:witold.suryn@etsmtl.ca)

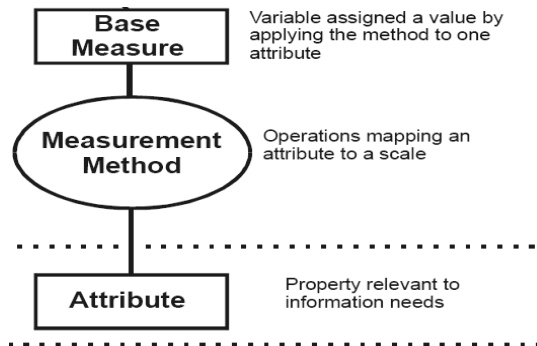
## Abstract

The ISO 9126 quality model is a 4-part suite of documents presenting 10 characteristics of the quality of software products, 27 sub characteristics, and an inventory of more than 250 derived measures proposed to quantify these quality characteristics and sub characteristics. However, these measures are presented only at a fairly abstract level as formulae composed from a set of 80 base measures. As the base measures themselves lack detailed descriptions, including the attributes they are attempting to measure, they are highly susceptible to individual interpretation. Improving the design of the 80 base measures would be a daunting task, not to mention the challenge of reaching an international consensus on all of them. The ISO 9126 standard is currently under revision by an ISO working group (ISO/IEC JTC1/SC7 WG6), and this paper proposes a process to determine which of these base measures should be improved in the timeliest fashion.

## 1. Introduction

The ISO 9126 quality model for software products is well known among researchers [1,2,3,4] and in the software industry [5,6,7]. The quality model in ISO 9126 has two sub models of software product quality (a shared sub model for internal and external quality and a separate sub model for quality-in-use), 10 quality characteristics, 27 sub characteristics, and more than 250 measures proposed to quantify these quality characteristics and sub characteristics. The 4-part suite of ISO 9126 is currently under revision by an ISO working group (ISO/IEC JTC1/SC7 WG6), and one of their challenges is to improve the definitions of these measures, which are proposed in Parts 2, 3, and 4 of ISO 9126. These parts have the status of ISO Technical Reports, since they are not yet considered mature enough to be recognized as International Standards. The 250+ measures are defined at a fairly high level as formulae built on a combination of ‘base measures’ and so are considered ‘derived’ measures, as defined in the ISO International Vocabulary on Metrology – VIM [8] – and in ISO 15939 [9]. While a derived measure corresponds to a combined set of base measures, every base measure

should correspond to a single, distinct software attribute. So, defining the attribute (e.g. the concept to be measured) should be the first step in defining a base measure (see Fig. 1), of which there are 80 in ISO 9126.



**Fig. 1 Relationship between an attribute and a base measure (ISO 15939)**

Each one of these attributes appears in one of the over 250 derived measures. For instance:

- the attribute ‘function’ appears in 38 derived measures, and therefore occurs in 15% of the derived measures;
- the attribute ‘duration’ appears in 26 derived measures, i.e. in 10% of them.

By contrast, a large number of attributes appear in a single derived measure.

However, as described in the 2003-2005 versions of ISO 9126, most of the attributes to be measured and their corresponding base measures are not documented at a detailed enough level to provide sufficient guidance to ensure the accuracy, repeatability, and repetitiveness of measurement results, in the event that the same software is measured by different measurers, which in turn leads to values that are potentially significantly different. To put it another way, while the numerical assignment rules for each derived measure are described as mathematical operations in the 2003-2005 versions of ISO 9126, neither the base measures for these operations nor the corresponding quality attributes have been described with sufficient clarity to ensure the quality of the measurement results.

Improving the design of these 80 base measures is a daunting task, considering the number of steps and iterations typically necessary to design software measures adequately, as illustrated in [10]. This design task is even more challenging when, in addition to the views of the person designing the measure, a consensus must be developed progressively at an international level, such as within an ISO committee composed of domain experts from a number of countries. Similarly, to determine which of these base measures must be improved in the timeliest fashion is a challenge. This paper proposes an approach (based on Pareto analysis from the Italian economist Vilfredo Pareto) to identify the priorities in sequencing the design of the base measures

needed for the next generation of ISO 9126 documents; that is, the upcoming ISO 25000 series, including the ISO 25021 technical report [11].

The rest of this paper is organized as follows. Section 2 presents an overview of the framework for designing software measures. Section 3 presents the suggested three-step approach to identify the priorities to be tackled in sequencing the measures to be designed, and which concepts are needed to carry out this task. Section 4 presents some examples of the application of this approach in the definition of some of the attributes to be measured and of corresponding base measures. Section 5 briefly discusses the applicability of the results obtained in the course of this research. Finally, section 6 constitutes a discussion of this issue, including comments about qualifiers to the definitions of some base measures.

## 2. A framework to define base measures

There are a number of hurdles to cross to improve the design of base measures, and of the definitions of their corresponding attributes for ISO 25021 [11]:

- First, with the exception of functional size measurement methods, the software engineering discipline is not supported by any other International Standard for software measures. As noted by Habra *et al.*, “In contrast to other fields of science and engineering, both researchers and practitioners must often design and develop their own individual software measurement methods, whereas these already exist in other fields of knowledge” [10].
- Second, some claim that, because software products are ‘intellectual products’, they cannot be measured. Habra *et al.* do not agree with this claim: “Although software products are most often viewed as intellectual artifacts, in a broader sense, they are also representations, through particular models, of physical phenomena inside a computer” [10].
- Third, very few measures in software engineering have been defined on the basis of a measurement principle, a measurement method, and measurement procedures.

The practical view of measurement presented by Habra *et al.* includes the three levels described in the ISO VIM (2004):

The **measurement principle** constitutes the scientific basis of measurement. For software entities (products), the measurement principle involves the model(s) used, and forms the basis on which to describe the entity for which a given attribute is intended to be measured. The idea is that modeling, as a central notion in software products, should be considered at the same level as scientific principles in other sciences and in engineering.

A **measurement method** is a generic operational description, i.e. a description of a logical sequence of operations, of the way to perform a measurement activity; that is, to move on from the attribute of an entity to be measured to the value representing the measurement result.

A measurement method should, in turn, be implemented by some concrete operations achieved through measuring instruments and/or practical operations: selection, counting, calculation, comparison, etc. This description of a measurement according to one or more measurement principles and to a given measurement method is called the **measurement procedure**. It is more specific, more detailed, and more closely related to the environment and to the measuring instruments (e.g. tools) than the method, which is more generic.

“The term measurement life cycle is used for the whole process of measurement involving the design of measurement method, the application of measurement method and the exploitation of the measurement results” [10]:. In terms of improving ISO 25021, the main interest is in the first phase, ‘the design of the measurement method’, which includes the following activities [10]:

- 1- Defining the measurement principle where this activity gives the precise description of what is going to be measured.
- 2- Defining a measurement method on the basis of that principle. This activity gives a general description of how to measure.
- 3- Determining an operational measurement procedure; that is, an implementation of the method in a particular context. This third activity gives a detailed description of how to measure [10].

It can be observed that, for most of its current base measures, the ISO 9126 standard neither provides the precise definition of the attribute (e.g. the ‘what’) that is being measured nor the generic description of how to measure it. Also, no operational measurement procedures are offered. For the upcoming version of ISO 25021, the ‘what’ and the ‘how’ should be spelled out by going through the above three steps to improve the design of the base measures.

Currently, only a very small number of the proposed base measures, such as those for the measurement of software functional size, have already gone through all these steps. Most do not even have a normalized definition of their attributes and therefore no precise description of what must be measured.

We have done an analysis for a subset of 31 attributes in ISO 9126. These were analyzed in terms of their maturity:

- first choice: no consensus or definition (19 attributes);
- the first activity has been performed, i.e. the definition of the base measure is in a standard (10 attributes);
- the second activity has been performed, i.e. a measurement method has been defined (1 attribute);
- the third activity has been performed, i.e. there is an operational measurement procedure (1 attribute);

### 3. Determination of priorities

Defining the full set of 80 attributes and necessary base measures is not considered a task which it will be feasible to perform within the next two to three years by the ISO 25021 editors' team.

#### 3.1 Analysis of occurrences

Table 1 presents a summary of the distribution of occurrences of the 80 distinct attributes to be quantified by a proposed base measure in ISO 9126–Part 2, 3, and 4:

- 47 of the 80 attributes appear in a single derived measure,
- 5 attributes have more than 10 occurrences,
- 15 attributes have from 3 to 10 occurrences,
- 13 attributes have 2 occurrences,
- 47 attributes have 1 occurrence, or 59% of the 80 different attributes

Occurrences of attributes in derived measures of ISO 9126-2, 3, and 4	Number of attributes	Percentage of attributes
1 occurrence	47	59%
2 occurrences	13	16%
From 3 to 10 occurrences	15	19%
More than 10 occurrences	5	6%
<b>TOTAL</b>	<b>80</b>	<b>100%</b>

**Table 1 Occurrences of attributes within the derived measures of ISO 9126**

#### 3.2 Analysis of the coverage of the derived measures

Table 2 presents an analysis of the coverage of the derived measures by each of the measured attributes, classified according to their number of occurrences. For example:

- The 47 attributes with 1 occurrence appear in 47 derived measures, that is, in less than 19% of the derived measures.
- The 5 attributes with more than 10 occurrences appear in 109 derived measures, that is, in 44% of the derived measures.

From Table 2, it can be seen that taking only the 20 attributes most frequently used (that is, with 3 or more occurrences), their corresponding set of 20 base measures would ensure a 71% coverage of the derived measures of the ISO 9126 quality model and with 13 more (that is, 33 attributes and corresponding base measures), or 81%; that is, an 81% coverage of the derived measures requires only the subset of 33 bases measures with 2 or more occurrences.

	<b>Attributes</b>	<b>Percentage of attributes</b>	<b>Derived measures</b>	<b>Percentage of derived measures</b>
Number of unique attributes	47	59%	47	19%
Attributes with 2 occurrences	13	16%	26	10%
From 3 to 10 occurrences	15	19%	69	28%
More than 10 occurrences	5	6%	109	43%
<b>Total</b>	<b>80</b>		<b>251</b>	<b>100%</b>

**Table 2 Coverage of the derived measures**

### **3.3 Coverage of the quality characteristics and sub characteristics**

In terms of the determination of priorities, it is important to verify whether or not all the characteristics and sub characteristics of the ISO 9126 quality models would be covered; that is, when defining these 33 attributes, is it possible to have at least one measure for every characteristic?

To answer this question, it is necessary to analyze the presence of attributes/base measures within the quality characteristics and sub characteristics. Table 3, column 2, shows the presence of at least one attribute/base measure for each characteristic in column 1.

<b>Characteristics</b>	<b>Attributes/base measures used</b>
<b>Internal/External</b>	
Functionality	9
Reliability	14
Usability	19
Efficiency	21
Maintainability	18
Portability	12
<b>Quality-in-use</b>	
Safety	3
Satisfaction	1
Productivity	3
Effectiveness	2

**Table 3 Occurrences of the 33 attributes in the 10 quality characteristics**

<b>Characteristics</b>	<b>Sub characteristics</b>	<b>Used</b>
Functionality	Accuracy	3
	Functionality compliance	1
	Interoperability	2
	Security	2
	Suitability	1
Reliability	Fault tolerance	3
	Maturity	5
	Recoverability	4
	Reliability compliance	2
Usability	Attractiveness	3
	Learnability	2
	Operability	9
	Understandability	4
	Usability compliance	1
Efficiency	Efficiency compliance	2
	Resource utilization	12
	Time behavior	7
Maintainability	Analyzability	5
	Changeability	6
	Testability	2
	Maintainability compliance	2
Portability	Stability	3
	Adaptability	2
	Coexistence	2
	Installability	2
	Portability compliance	3
	Replaceability	3

**Table 4 Occurrences of the 33 attributes in the 27 quality sub characteristics**

## 4. Definition of the attributes

### 4.1 Attributes with more than 10 occurrences

There are 5 attributes – function (38), duration (26), task (18), case (16), and failure (11) – which appear between 11 and 38 times.

1- The attribute ‘function’ is consistently used with ‘number of...’ in ISO 9126–Parts 2 to 4. However, nowhere is it defined precisely, and its interpretation in practice can vary considerably across individuals, technology, functional domains, etc. Notwithstanding this, in ISO 9126, the industry has developed various consensus over the years on the measurement of the functional size of software. This has led to the adoption of 5 international standards for functional size measurement that could also be used as normalization factors in quality measurement, such as in the measurement of defect density.

2- The attribute ‘duration’ is a length of time in seconds, minutes, hours, etc. The ‘second’ as a unit of measurement is already well defined and is a part of the set of international standards for units of measurement.

3- The attribute ‘task’ has multiple definitions within ISO standards:

- a sequence of instructions treated as a basic unit of work by the supervisory program of an operating system, in ISO 24765 [13].
- in software design, a software component that can operate in parallel with other software components, in ISO 24765 [13].
- the activities required to achieve a goal, in 4.3 of ISO TR 9126-4 [4].
- a concurrent object with its own thread of control, in ISO 24765 [13].
- a term for work, the meaning of which and placement within a structured plan for project work varies by the application area, industry, and brand of project management software, in the PMBOK in ISO 24765 [14].
- required, recommended, or permissible action, intended to contribute to the achievement of one or more outcomes of a process, in section 4.5 of ISO 12207 [15] and in section 4.34 of ISO 15288 [16].

Therefore, for ‘task’, it is necessary to revise each usage of task for each attribute in each quality characteristic and sub characteristic.

4- The attribute ‘case’ (with 16 occurrences) is not defined in the ISO 9126 standard, but is defined as follows in ISO 24765[13]:

*“a single-entry, single-exit multiple-way branch that defines a control expression, specifies the processing to be performed for each value of the control expression, and returns control in all instances to the statement immediately following the overall construct.”*



5- The attribute 'failure' is quite challenging, since it has multiple definitions:

- termination of the ability of a product to perform a required function or its inability to perform within previously specified limits – see 4.2 in ISO 25000 [17].
- the inability of a system or component to perform its required functions within specified performance requirements – ISO /IEC 24765 [13].
- an event in which a system or system component does not perform a required function within specified limits – IEEE 982.1 [18].
- the termination of the ability of a functional unit to perform its required function – IEEE 982.1 [18].

NOTE: A failure may be produced when a fault is encountered.

The first definition of the attribute 'failure' could be suggested, but should be revised in the context of each attribute in each quality characteristic and sub characteristic.

## 4.2 Attributes with more than 3 occurrences

There are 15 attributes with more than 3 occurrences. The same type of analysis can be performed with those attributes.

The attributes 'fault', 'questionnaire', 'size', 'warning messages', and 'error' are already defined in ISO or IEEE standards. The attributes 'compliance item', 'item', and 'data item' have the word 'item' in common. This means that a good definition of 'item' would be very useful.

It should be noted that, even if a number of definitions exist in the standards (or in the literature) for the terms mentioned above, this does not mean that it is necessarily recommended to use them as is:

- these definitions might not have been tested operationally,
- the definition might not be useful for a measurement context.

## 4.3. The attribute qualifiers

In addition to the 80 different base measures, ISO 9126–Parts 2, 3, and 4 include a number of qualifiers which characterize some aspects of the base measures.

- For example, the base measure 'number of failures' may refer at times to the number of **resolved** failures or to the number of failures **actually detected**. The terms 'resolved' and 'actually detected' are referred to here as qualifiers of the term 'failures' for this base measure; that is, they qualify a subset of the same attribute.

Sometimes the qualification of the base measure uses a broader qualifier.

- For example, the number of ‘critical and serious failure occurrences avoided’.

The qualifiers in the ISO 9126 quality model are, most of the time, added to measures using a sentence, not just a word. A solution would be to suggest, whenever possible, a reference in the set of ISO standards. For example, ‘type of maintenance’ could be aligned, along with its corresponding concepts, with the ISO standard on software maintenance, that is, ISO 14764 [12].

Another possibility, when there is no reference to a standard for specific qualifiers, would be to modify them when relevant. To have defined the important attributes for the ISO 9126 quality model is an important improvement. After completing the priorities in this first iteration of improvements, further research will be necessary to precisely define the ‘qualifiers’.

## **5. Considerations on the applicability of the research results obtained**

As has been illustrated throughout this paper, the base measures and their current lack of adherence to metrology principles and characteristics constitute one of the major impediments to the overall applicability of ISO 9126. It is also recognized that, to properly analyze, verify, and correct all the 80 base measures would require considerable time and effort on the part of a single, isolated research team, which could make the results obsolete before they became publicly available. This provided the motivation for setting up a larger, multi-group research team and program to work concurrently, including the research groups at the École de technologie supérieure (ÉTS) – Université du Québec (Canada) and researchers at the Middle East Technical University – METU (Turkey), and within the work in progress of ISO/IEC JTC1/SC7 Working Group 6. The plan is to subdivide the original program into smaller projects which will be assigned as Master’s degree thesis topics to students at both universities, with the participation of ISO/IEC JTC1/SC7 experts on software measurement. Organizing the work in this way will facilitate the transition of the research results to the ISO normative level and, consequently, enable quicker integration of adequately structured base measures into the ISO 25000 series of standards for the measurement of the quality of software products.

## 6. Summary

The over 250 derived measures of the ISO 9126 standard are described at a fairly abstract level as formulae composed of a set of 80 base measures. As the base measures themselves lack detailed descriptions, including well characterized definitions of the attributes they are attempting to measure, there is too much scope for interpretation in individual measurements.

Improving the design of these 80 base measures is a daunting task, not to mention the need to reach an international consensus on all of them. This paper has proposed a process to determine which of the base measures must be improved in the timeliest fashion, considering the various hurdles that must be overcome.

In particular, improvement work should focus on 5 of the 80 attributes in ISO 9126 that is, function (38 occurrences in derived measures), duration (26), task (18), case (16), and failure (11). Work on the detailed design of the base measures and on the definitions of the attributes should leverage relevant measurement definitions from other international standards wherever possible. Even definitions from existing standards still need further refinement to facilitate their use in operational procedures from a measurement viewpoint. Finally, the ISO/IEC 9126 standard also includes a number of qualifiers to the base measures which will require further clarification from a measurement viewpoint.

In conclusion, much work remains to be done to define the base measures in detail, even those identified as requiring priority attention. The SC7 WG6 team should probably consider completing only one or two activities in the first phase, the design of the measurement method, for a number of them.

### References

1. ISO, 2001, ISO/IEC IS 9126-1, Software Engineering – Product Quality – Part 1: Quality model, International Organization for Standardization, Geneva (Switzerland).
2. ISO, 2003, ISO/IEC TR 9126-2, Software Engineering – Product Quality – Part 2: External Metrics, International Organization for Standardization, Geneva (Switzerland).
3. ISO, 2003, ISO/IEC TR 9126-3, Software Engineering – Product Quality – Part 3: Internal Metrics, International Organization for Standardization, Geneva (Switzerland).
4. ISO, 2004, ISO/IEC TR 9126-4, Software Engineering – Product Quality – Part 4: Quality-in-Use Metrics, International Organization for Standardization, Geneva (Switzerland).

5. Abran, A., Al-Qutaish, Rafa E., Desharnais, J. M., Habra, N., An Information Model for Software Quality Measurement with ISO Standards, SWEDC-REK – International Conference on Software Development, Reykjavik, Iceland, University of Iceland, 2005, pp. 104-116
6. Al-Qutaish, Rafa E., SPQMM: A Software Product Quality Maturity Model using ISO/IEEE Standards, Metrology, and Sigma Concepts, PhD Thesis, Department of Software Engineering and IT, École de technologie supérieure- Université du Québec, Montréal, Canada, 2007
7. Abran, A., Al-Qutaish, Rafa E., Desharnais, J. M., Harmonization Issues in the Updating of the ISO Standards on Software Product Quality, Metrics News ,vol. 10, no. 2, December, 2005, pp. 35-44
8. ISO VIM, 2004. International Vocabulary of Basic and General Terms in Metrology (Draft). International Organization for Standardization – ISO Geneva.
9. ISO, 2002, ISO/IEC IS 15939, Software Engineering – Software Measurement Process, Geneva: International Organization for Standardization.
10. Habra, N., Abran, A., Lopez, M., Sellami, A., A Framework for the Design and Verification of Software Measurement Methods, Journal of Systems and Software, Elsevier, vol. 81, issue no.5, May 2008, pp. 633-648.
11. ISO, 2008, ISO/IEC TR 25021, Software Engineering: Software product Quality Requirements and Evaluation (SQuaRE) – Quality measure elements, International Organization for Standardization, Geneva (Switzerland).
12. ISO/IEC 14764:2006, Software Engineering – Software Life Cycle Processes – Maintenance, International Organization for Standardization, Geneva (Switzerland).
13. ISO 24765, Systems and Software Engineering Vocabulary, International Organization for Standardization, Geneva (Switzerland).
14. PMBOK, ‘A Guide to the Project Management Body of Knowledge (PMBOK® Guide),’ Project Management Institute, Third Edition
15. ISO 12207:2008 Systems and software engineering – Software life cycle processes, International Organization for Standardization, Geneva (Switzerland).
16. ISO 15288:2008 Systems and software engineering – System life cycle processes. International Organization for Standardization, Geneva (Switzerland).
17. ISO 25000:2005 Software Engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Guide to SQuaRE, International Organization for Standardization, Geneva (Switzerland).
18. IEEE 982.1-1988 IEEE Standard Dictionary of Measures to Produce Reliable Software.