Software Metrics & Software Metrology

Alain Abran

Chapter 4
Quantification and Measurement are Not the Same!

Agenda

This chapter covers:

- The difference between a number & an analysis model.
- The Measurement Information Model in ISO 15939:
  - its metrology-related perspective
  - its analysis perspective for the quantification of relationships.
  - Examples of these differences within a Measurement Information Model.
- A Metrology Design ≠ A Quantification Model of Relationships
  - Example 1: the measurement of a single attribute
  - Example 2: the quantification of relationships across attributes and entities.

Agenda

This chapter covers:

- **The <u>difference between a number & an analysis model</u>**
- The Measurement Information Model in ISO 15939:
  - its metrology-related perspective
  - its analysis perspective for the quantification of relationships.
  - Examples of these differences within a Measurement Information Model.
- A Metrology Design ≠ A Quantification Model of Relationships
  - Example 1: the measurement of a single attribute
  - Examples 2: the quantification of relationships across attributes and entities.

© 2010 Alain Abran - Software Metrics & Software Metrology

Introduction: Numbers are not all created equals

- Software practitioners and researchers alike often forget that numbers are not all created equal:

  - A number derived from the result of a measurement process which meets the metrology requirements [1] is a quantity expressed with a measurement unit.
    - A number with a measurement unit obtained through the proper application (manual or automatic) of its corresponding measurement method will have many more measurement qualities (in the metrology sense) than a number derived from opinion only.

  - A number derived from a mix of mathematical operations without consideration of measurement units and scale types will still be a number, but it could be a meaningless one.
    - Practitioners may feel good about models which appear to take into account a large number of factors (i.e. as in many estimation models and quality models). However, feeling good does not add validity to mathematical operations that are inadmissible in measurement.
      - For example: some of the Halstead's metrics – see chapter 7.
      - For example: see the Use Case Points - see chapter 9.

[1] See chapter 3 on metrology.

Introduction: Numbers are not all created equals

- In practice, various types of quantitative models produce numbers in outputs (i.e. the outcomes of the models) which do not have the same qualities as numbers which meet the requirements of metrology:
  - An estimation model will provide a number as an estimate:
    - to every such estimated number is associated a (potentially large) range of variations, depending on the number of input parameters and their corresponding uncertainties, as well as on the uncertainties about the relationships across all such input parameters.
    - <u>these estimated numbers are not meaningful without a knowledge (& understanding) of the corresponding uncertainties</u>.
  - A quality model will provide a number which typically depends on:
    - a specific selection among a (potentially large) number of alternatives,
    - the assignment of a percentage to each contributing alternative, which assignment is based on the opinion of one person (or a group of persons) and
    - comparison of each contributing alternative with distinct threshold values, which themselves are often defined by opinion as well.

Introduction: Numbers are not all created equals

- In many instances, in these analysis models
  - some, if not all, of the numbers used as inputs to these models are obtained by opinion, rather than from precise measurements (with measurement instruments or from the application of detailed measurement procedures);
  - these numbers are combined without explicitly describing the admissible mathematical operations and treatment of the corresponding measurement units; and
  - the outcomes of such models are indeed numbers, but they do not have metrological properties [1], and should be handled very cautiously.
- Analysis models like these are quantitative models, but they are not measurement models in the metrological sense.
  - Such differences between quantitative analysis and measurement are not generally discussed in the software engineering literature.

[1] See chapter 3.

ISO 15939 Measurement Information Model

- The Measurement Information Model from ISO 19539 (Next Slide) sets out the various steps necessary for the design an *information product* when a measurable concept has to be designed and used in practice.
- From the bottom up:
  1. A specific *measurement method* has to be designed to obtain a *base measure* for a specific attribute.
  2. The values of 2 or more base measures can be used next in a computational formula (by means of a measurement function) to construct a specific *derived measure*.
  3. These derived measures are used next in the context of an *analysis model* of relationships to construct an *indicator*.
  4. Then, the indicator (i.e. the number from point 3 above) is used for interpretation purposes to build the information product to meet the information needs.
     - This means that the indicator's value is interpreted within the prescribed context as describing, in the language of the measurement user, an information product for his information needs [ISO 15939].

ISO 15939 Measurement Information Model

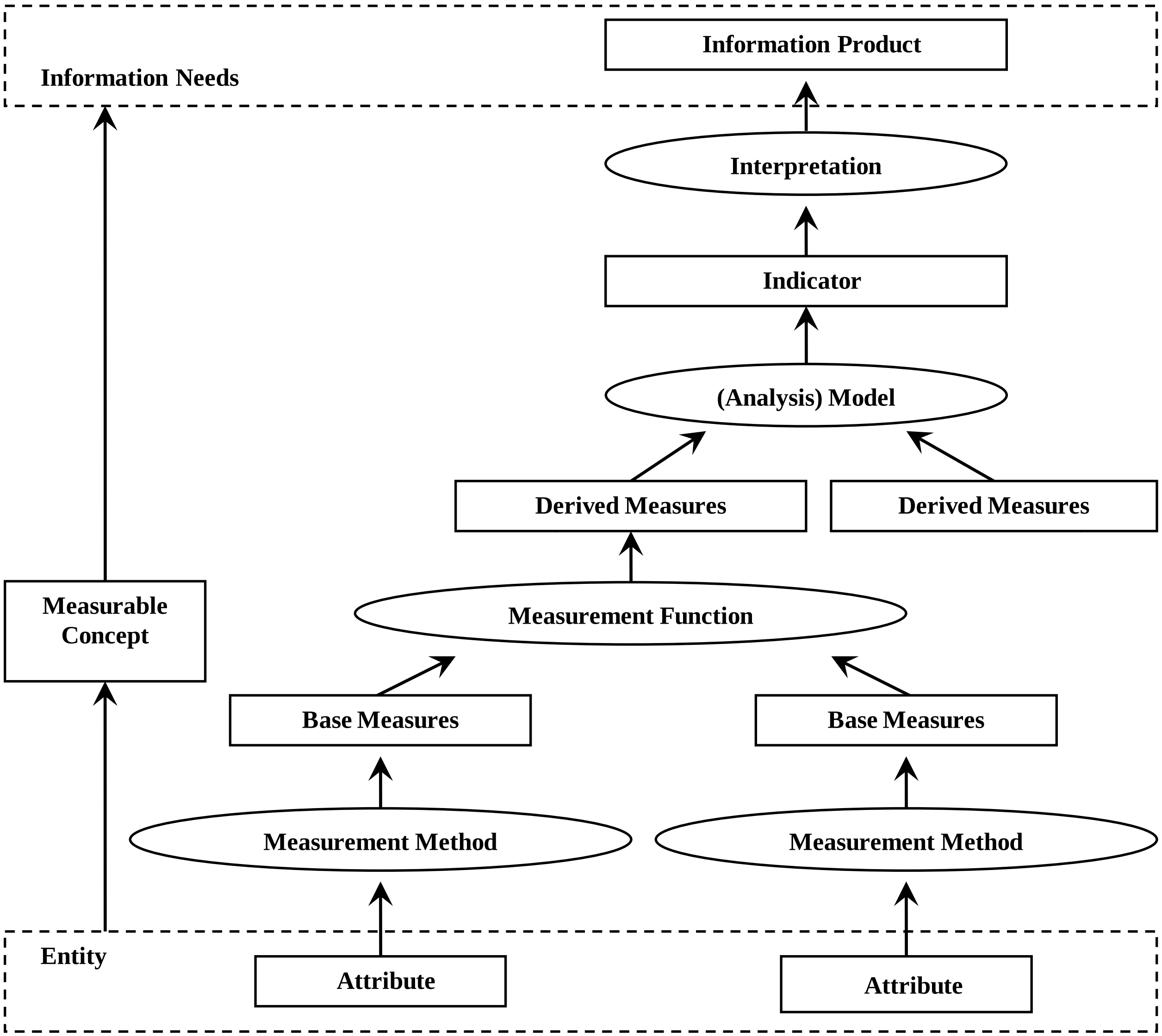


Figure: Measurement Information Model from ISO 15939

© 2010 Alain Abran - Software Metrics & Software Metrology

ISO 15939 Measurement Information Model

NOTE:

- The derived measures & the indicator inherit the properties of the mathematical operations on which they are built:

  - These numbers are meaningful:
    - when derived from admissible mathematical operations.

  - These numbers are meaningless:
    - when derived from inadmissible mathematical operations, or
    - when the measurement units and measurement scale types are not considered correctly within the mathematical operations.

ISO 15939 Measurement Information Model

- The direction of the arrows in previous Figure has been added to the ISO 15939 Measurement Information Model to highlight the sequence of steps required to implement an already well defined Information Product:
  - i.e. from the detailed measurement of the base measures up to the interpretation of the information, that is, the end Information Product.

    - Example: When an organization already collects measures of project effort and software size, it uses then such measures to:
      - build its own productivity models of past projects and next,
      - uses such models to prepare estimates for its selection of future projects.

ISO 15939 Measurement Information Model

- When the Information Product structure does not yet exist, an organization would typically work in top-down fashion, by
  - starting with specification of the information product it needs, and
  - working top-down to define the detailed analysis and measurement processes required to fulfill its information needs.

ISO 15939 Definitions Adapted from the VIM

The terminology of the ISO International Vocabulary of Basic and General terms in Metrology (VIM) has been adopted in ISO 15939 as the agreed-upon measurement terminology for software and system engineering-related ISO standards (with a few adaptations to facilitate its use within the software engineering community, which was previously using a variety of non standardized terms).

The VIM defines *base quantity* and *derived quantity*, while ISO 15939 uses the expressions *base measure* and *derived measure* for the corresponding VIM concepts.

This adaptation of terms was designed to facilitate the adoption of ISO 15939 in the software engineering community. See also the measurement definitions in Appendix B.

Scope of the ISO 15939 Measurement Information Model

To better understand the Measurement Information Model in ISO 15939, it is useful to identify what in the model is related to metrology concepts and what is not.

A- Metrology-related: measurement of the attribute of an entity

- The bottom portion of the ISO 15939 Measurement Information Model can be mapped to the metrology concepts in 2 steps – see Figure in next slide:
  1. **Data Collection**: when a measurement method is used to measure an attribute [1], the corresponding output is the base measure of the specific entity being measured.
     - This corresponds to the **data collection** of the base measure for each entity being measured;
  2. **Data Preparation**: when a number of the base measures of the data collected are combined through a measurement function (using agreed-upon mathematical formula and related labels), then the combined units are considered as derived measures.
     - This corresponds to **data preparation**, prior to the analysis phase.

[1] Of course, the attribute must be well defined; if not, it is pretty challenging to design an adequate measurement method.

Scope of the ISO 15939 Measurement Information Model

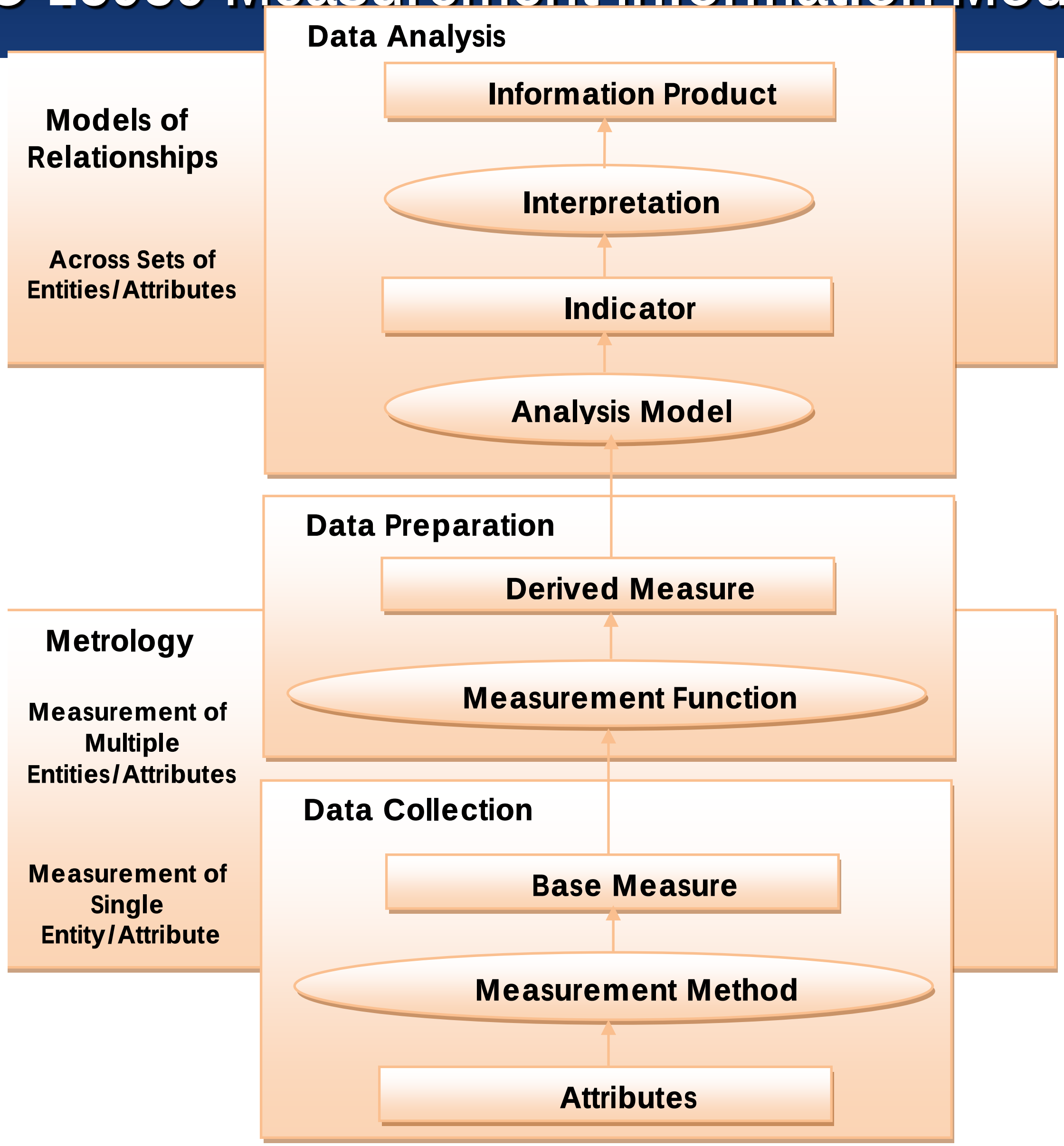


Figure: The 2 perspectives of the ISO 15939 Measurement Information Model

Scope of the ISO 15939 Measurement Information Model

B- Non Metrology-related: quantification of relationships across attributes and entities

- The top portion of the ISO 15939 Measurement Information Model deals with the analysis (through quantification) of relationships across entities and attributes.
- This analysis part of the ISO 15939 Measurement Information Model includes:
  1. Analysis model: e.g the modeling of the relationships across entities and attributes [1] to derive an indicator of the value of such relationships.
  2. Interpretation: the indicator would then be interpreted to produce the Information product that would typically be used next in an evaluation or decision making process.

[1] See section 5 and Figure 4 of this chapter for more details.

Scope of the ISO 15939 Measurement Information Model

- The metrology-related bottom part of the Measurement Information Model is supported by the set of metrology concepts, as described in Chapter 3.

- The upper part of the Measurement Information Model is outside the scope of the VIM, since it deals with the use of the measurement results from the lower part of the model.
  - This analysis is not extensively described in ISO 19539, except through a few complex definitions tailored to that specific standard

- These 2 perspectives are discussed in more details next.

Agenda

This chapter covers:

- The difference between a number and an analysis model.
- The Measurement Information Model in ISO 15939:
  - **its metrology-related perspective**
  - its analysis perspective for the quantification of relationships.
  - Examples of these differences within a Measurement Information Model.
- A Metrology Design ≠ A Quantification Model of Relationships
  - Example 1: the measurement of a single attribute
  - Example 2: the quantification of relationships across attributes and entities.

The Metrology Perspective in the ISO 15939 Measurement Information Model

Metrology-related perspective of the ISO 15939 Measurement Information Model – i.e. the bottom part of previous Figure.

- **Data collection: base measures**
  - Every base measure must correspond to a single, distinct, software attribute:
    - i.e. a property of an object or concept
    - So, identifying the attribute of the entity to be measured and quantifying it through its measurement method corresponds to the data collection step – see next slide figure
- See section 3.3 in Chapter 3 for the VIM definitions of base quantities & derived quantities.

The Metrology Perspective in the ISO 15939 Measurement Information Model

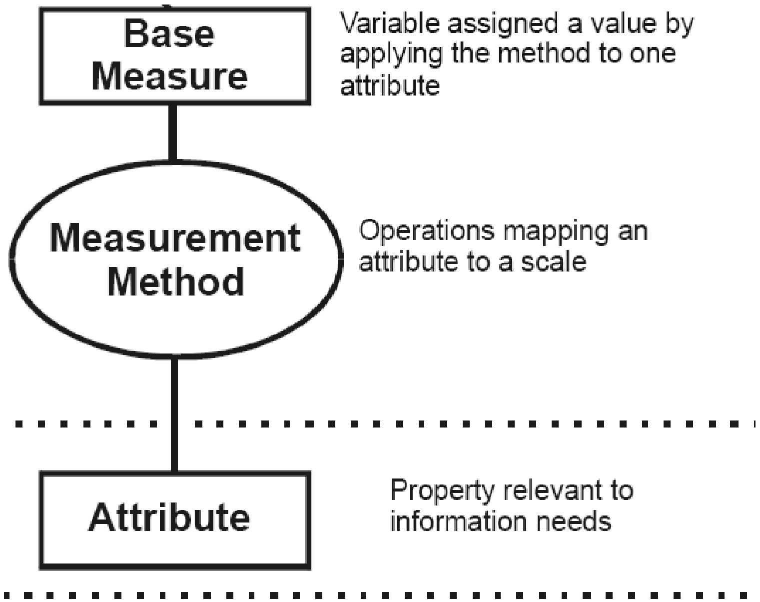


Figure: The data collection of a base measure for an attribute (ISO 15939)

The Metrology Perspective in the ISO 15939 Measurement Information Model

- **Data preparation: derived measures**
  - Depending on the Information Needs, some of the base measures already collected for an entity can be assembled according to a measurement function (e.g. a computational formula) defined for each derived measure – see Figure: the Data Preparation step:
    - A derived measure is therefore the product of a set of measurement units properly combined (through a measurement function).

Addition of the same units ≠ Derived Measure

The addition of houses (of the same types or of different types) gives a total number of 'houses': the addition of base measures of the same type does not lead to derived measures.

Similarly, the additions of 'numbers of points' (however 'points' have been defined) lead to a total number of 'points': this does not represent a derived measure.

-

[1] See section 3.3 in Chapter 3 for the VIM definitions of base quantities & derived quantities.

The Metrology Perspective in the ISO 15939 Measurement Information Model

- If a derived measure is designed bottom-up, the name assigned to this combination of units should correspond to the concept representing the particular combination of measurable attributes.
  - The accuracy of a derived measure (together with the corresponding measurement errors) is directly related to:
    - the accuracy of each of its base measures, and
    - how these base measures are mathematically combined [1].
  - Stated differently: the qualities of the corresponding measuring device(s) of the base measures impact the quality of the derived measures.

[1] See chapter 5.

The Metrology Perspective in the ISO 15939 Measurement Information Model

Example: Accuracy depends on base measures

The accuracy of a measurement of velocity will directly depend on the accuracy of its 2 base measures: distance and time.

- When their corresponding base measures are not sufficiently well defined, standardized, and instrumented to ensure the accuracy, repeatability, and repetitiveness of measurement results, then, when the same entity (software) is measured by different measurers, the results can potentially be significantly different.
  - Note: a derived measure is *descriptive*. It does not explain a relationship, nor does it say anything about the strength of such a relationship across distinct attributes.

The Metrology Perspective in the ISO 15939 Measurement Information Model

Example of a Derived Measure: Velocity

The combination *distance traveled over a period of time* (e.g. km/hour) is associated with the concept of velocity.

Such a derived measure (i.e. velocity) is typically measured by a measuring device which:

– captures both base measures simultaneously (that is, distance and time, measured in meters and seconds on a car's speedometer, for instance),

– has an integrating feature which divides the base measures to produce a ratio (time/distance) to represent the velocity concept, and

– has a display feature which shows up the measurement results using a standardized display convention:

For example, converting meters per second into the universally adopted standard for cars, which is 'kilometers per hour'.

It must be observed that the result of the mathematical operations must also lead to the combination of the measurement units of its corresponding base measures.

The Productivity Ratio: A Derived Quantity

The ratio of the outputs of a process to the inputs to this specific process (such as the number of cars produced by 1,000 work-days, or number of function points per person-month) is associated with the productivity concept.

This productivity ratio is entirely descriptive: it does not attempt to express why a given production process has such a value (i.e. such a ratio).

Agenda

This chapter covers:

- The difference between a number and an analysis model.
- The Measurement Information Model in ISO 15939:
  - its metrology-related perspective
  - its analysis perspective for **the quantification of relationships**.
  - Examples of these differences within a Measurement Information Model.
- A Metrology Design ≠ A Quantification Model of Relationships
  - Example 1: the measurement of a single attribute
  - Example 2: the quantification of relationships across attributes and entities.

The Quantification of Relationships in ISO 15939

- ***Quantitative Elements of the ISO 15939 Analysis Model***
  - The top part of the ISO 15939 Measurement Information Model deals with the 3rd step of the Measurement Context Model presented in Chapter 2:
    - <u>the use of measurement results</u> in various evaluation or decision making models.
  - This use of measurement results is represented very succinctly in ISO 15939 with:
    - Analysis Model
    - Interpretation
    - one number: Indicator
  - In practice, however, this use of measurement results typically involves:
    - **analysis of the relationships across different measurement results** with respect to various conditions within a context, and
    - **assessment against reference contexts** for evaluation and/or decision making.

The Quantification of Relationships in ISO 15939

- A number of concepts within these descriptions do not appear in the Measurement Information Model of ISO 15939, such as:
  - Decision criteria
  - Assumptions
  - Expected relationships
  - Estimates or evaluation
  - Numerical thresholds or targets
  - Statistical confidence limits, etc.

The Quantification of Relationships in ISO 15939

The ISO 15939 Definitions for the Use of Measurements Results

Indicator
An indicator is a measure providing an estimate or evaluation of specified attributes derived from a model with respect to defined information needs. Indicators are the basis for analysis and decision making. These are what should be presented to measurement users.

(Analysis) Model
An algorithm or calculation combining one or more base and/or derived measures with associated decision criteria.
It is based on an understanding of, or assumptions about, the expected relationship between the component measures and/or their behavior over time. Models produce estimates or evaluations relevant to defined information needs.
The scale and measurement method affect the choice of analysis techniques or models used to produce indicators.

Decision criteria
Decision criteria are numerical thresholds or targets used to determine the need for action or further investigation, or to describe the level of confidence in a given result. Decision criteria help to interpret the results of measurement. Decision criteria may be calculated or based on a conceptual understanding of expected behavior.
Decision criteria may be derived from historical data, plans, and heuristics, or computed as statistical control limits or statistical confidence limits.

NOTE: Some of the terms not represented in the model in Figure 1 are underlined above in the descriptions.

The Quantification of Relationships in ISO 15939

- ***Refined representation of the Analysis Model***
  - Figure 4 includes 2 additional major blocks – see next figure:

1. <u>A standard reference model (Figure 4, bottom left)</u>, which can include, for instance, an accepted model of the relationships across distinct types of objects of interest. When such a reference model exists, this can be:
   - an industry model
   - an ISO model
   - a generally accepted statistical technique (and related mathematical model).

- This standard reference model would include:
  - the set of formal (or informal and assumed) individual relationships, together with the base or derived measures to be considered as evaluation or decision criteria
  - the algorithm (mathematical or implied) that combines them in an (implied) criterion.

Chapter 10 presents an illustration of a standard reference context represented by the ISO 9126 quality models.

The Quantification of Relationships in ISO 15939

2. An organizational reference context (Next Slide Figure, upper left), ideally aligned with the standard reference, with a set of selection criteria and values specific to the organization:
   - this organizational reference would contain the reference values necessary for interpretation:
     - a set of reference values specified for this context
     - evaluation or decision criteria with either:
       - target values, or
       - specific evaluation scales

The Quantification

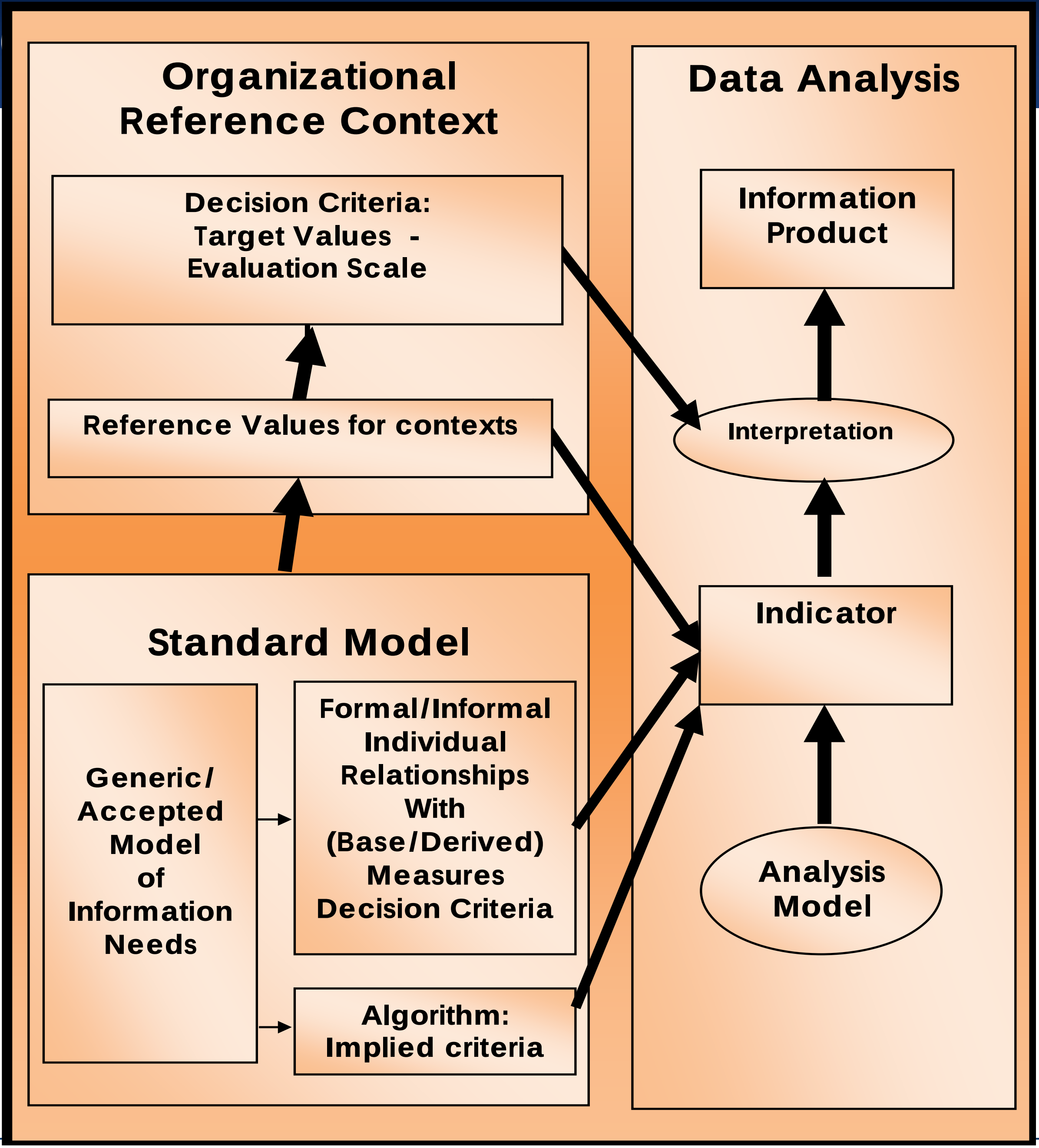


Figure 4 : Refined Analysis Model of the ISO 15939 Measurement Information Model

The Quantification of Relationships in ISO 15939

- ***The (Implicit) Link between Measurement and Quantitative Analysis of Relationships***
  - In the Measurement Information Model of ISO 15939, the link between the 2 major parts (that is, the Metrology-related bottom part of measurement and the Analysis-related upper part of quantification) are not explicitly described:
    - ISO 15939 makes the assumption that this link exists and that it is complete on its own.
    - In practice, the issue is more complex, in particular in domains where measurement and quantification (or either one) are not yet mature.

Agenda

This chapter covers:

- The difference between a number and an analysis model
- The Measurement Information Model in ISO 15939:
  - its metrology-related perspective
  - its analysis perspective for the quantification of relationships.
  - <u>Examples of these **differences with a productivity ratio & a productivity model**</u>
- A Metrology Design ≠ A Quantification Model of Relationships
  - Example 1: the measurement of a single attribute
  - Example 2: the quantification of relationships across attributes and entities.

A Productivity Model: An ISO 15939 Measurement Information Model

- The next slides illustrate the differences between the metrology and non-metrology related part of the ISO 15939 Measurement Information Model by looking at the differences between a productivity ratio and a productivity model.
  - Ex: statistical techniques propose ‘standard reference models’ to facilitate the analysis of relationships embedded in software productivity and estimation models.

A Productivity Model: An ISO 15939 Measurement Information Model

- ***A productivity model is more than a productivity ratio***
  - There are major differences between a productivity ratio and a productivity model.
    - A productivity ratio is related to the metrology-related part of the ISO 15939 Measurement Information Model: the productivity ratio is strictly defined as composed of 2 base measures (Output over Input):
      - This productivity ratio is based strictly on the measurement, from a metrology perspective, of the respective distinct attribute of the corresponding 2 distinct entities representing the output and the input (for example: Output = Function Points and Input = work-hours).
      - This productivity ratio is strictly descriptive and limited to what is being measured.

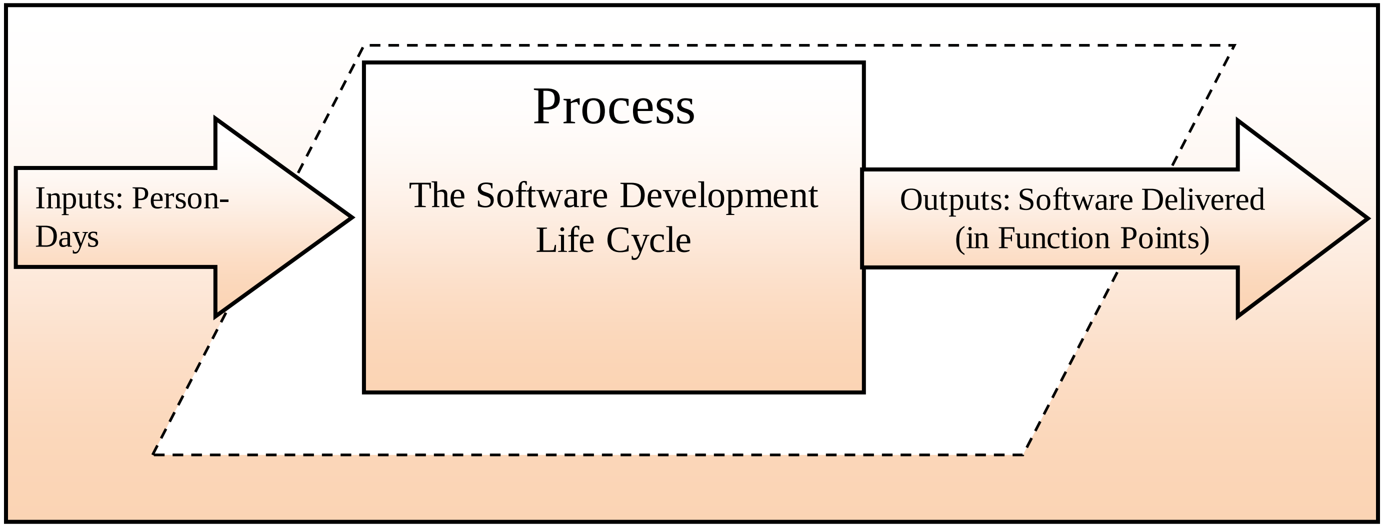
A Productivity Model: An ISO 15939 Measurement Information Model

- If we move now from a productivity ratio to a productivity model, a number of additional elements are added, since the purpose becomes:
  - the analysis of relationships across many entities (e.g. many projects) that have been measured, and often
  - an estimation of what would happen should this production process (which has been quantified indirectly through the measurement of its output and input) be used again to estimate the next project.

- To explore this, let us look at the next slide, which illustrates a production process in its simplest form:
  - The Input is on the left.
  - The Process is in the middle.
  - The Output is on the right.

A Productivity Model: An ISO 15939 Measurement Information Model

Figure 5: A Software Production Process



A Productivity Model: An ISO 15939 Measurement Information Model

Note:

- the productivity ratio has 2 explicit dimensions (Input and Output) that are explicitly present in a productivity ratio,
- the production model has another implicit dimension as well, that is, the production process itself.

- The objective is now **to quantify** this production model, typically using a productivity model (more commonly referred to as an estimation model in software engineering) rather than to measure it.

A Productivity Model: An ISO 15939 Measurement Information Model

A quantitative representation of a production process is typically built by:

- collecting the base measures of the production process over a number of completed projects, for example:
  - Input = Effort (in work-hours or work-days)
    - See next slide for the information required to ensure consistency in the measurement of the 'effort' variable in a multi-organizational data repository

      Similar rules must of course apply within a single organization to ensure that the effort recorded is recorded consistently across individuals and work groups
  - Output = Functional Size of the completed software (in Function Points)
- quantitatively modeling the relationships across these 2 variables, where:
  - effort is the dependent variable, and
  - functional size is the independent variable.

A Productivity Model: An ISO 15939 Measurement Information Model

- Statistical techniques can be considered as **standard reference models** for modeling these relationships
  - each statistical technique with a distinct mathematical representation (and corresponding strengths, constraints, and limitations).

- The productivity model, in its simplest form with a single dependent variable (x), could be expressed as:

$$x = f(y),$$

Where:

- the <u>dependent</u> variable 'x' would be in work-hours
- the <u>independent</u> variable 'y' would represent the size of the software (in Function Points).

© 2010 Alain Abran - Software Metrics & Software Metrology

A Productivity Model: An ISO 15939 Measurement Information Model

- The quantification of the productivity ratio:
  - is the outcome of the measurement of 2 entities (that is: Inputs & Outputs)
  - while the meaning of the division of these 2 numbers represents something different: the performance (in the sense of productivity) of a third entity, the process itself.
    - This means that the measurement of the productivity of the process is derived not from a direct measurement of the process, but from an indirect measurement of 2 other entities (the Inputs & Outputs of the process).

- The Analysis Model considers a number of distinct dimensions and combine them, in some manner, into a single number.
  - This corresponds to the definition in ISO 15939 that a model is an algorithm combining one or more base and/or derived measures, along with their associated decision criteria.

A Productivity Model: An ISO 15939 Measurement Information Model

Note:

- while a derived measure gives a combination of units
  - e.g. Function Points per work-hour

- the productivity model produces as output a single quantity with its corresponding single unit of the dependent variable:
  - the output of the F(x) analysis model is strictly in work-hours (e.g. Effort)

    (even though many additional independent variables could have been taken into account in more comprehensive productivity models - i.e. estimation models).

Next: 2 examples of productivity models based on statistical techniques:

- An average
- A linear regression

A Productivity Model: An ISO 15939 Measurement Information Model

- **Productivity model built with an averaged productivity**
  - The average statistical function can be considered as the algorithm of a standard reference model.
    - A productivity model built using an averaged productivity is presented in Next Slide.
  - An average is a well-known mathematical function, with corresponding properties (and limitations as well).
  - This average productivity is built by:
    - calculating the productivity ratios of each single project within a sample, then
    - adding them up, and
    - dividing by th.e number of projects in the sample

A Productivity Model: An ISO 15939 Measurement Information Model

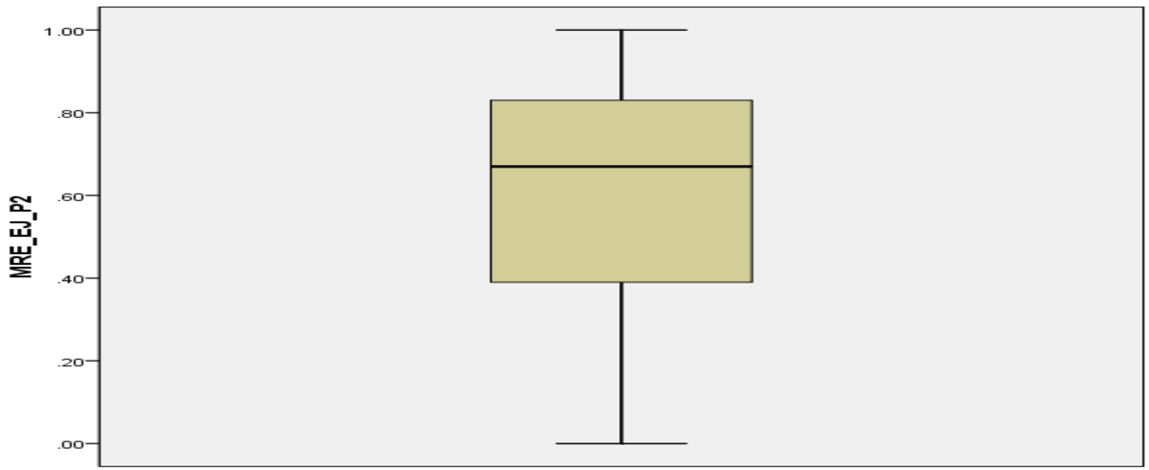


Figure 6: An averaged productivity and quartile box-plot

© 2010 Alain Abran - Software Metrics & Software Metrology

A Productivity Model: An ISO 15939 Measurement Information Model

- Note: this average describes the full sample, and not the individual projects in the sample.
  - In addition, a number of related characteristics are typically provided with the standard average function, such as:
    - minimum
    - maximum
    - first quartile, last quartile
    - 1-standard deviation
    - 2-standard deviation, etc.
    - skewness
    - kurtosis
    - etc.
- Some of these are represented graphically in the previous slide, which presents both the average of a sample (i.e. the horizontal line within the grey box), as well as the box-plot of quartiles.

A productivity model built with a linear regression.

- The linear regression statistical function – see below - can be considered as the algorithm of a standard reference model.

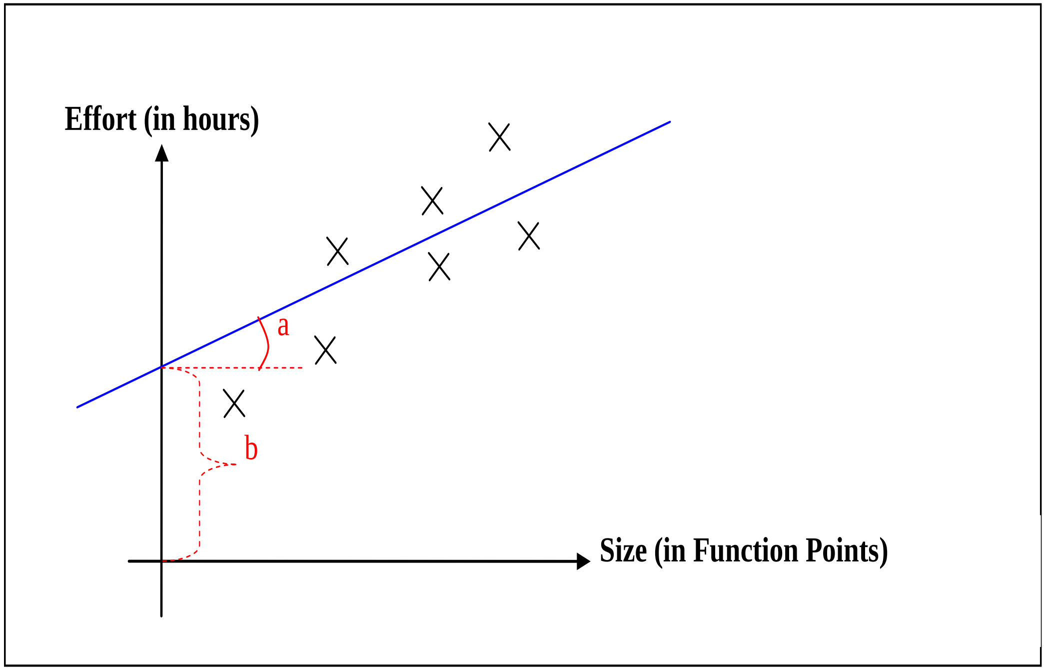


Figure 7: A Production Model with variable costs a) and fixed costs b)

A Productivity Model: An ISO 15939 Measurement Information Model

- The quantitative representation from the linear regression statistical technique is of the following form:
  - The dependent variable of Effort is a function of the independent variable of functional Size, that is: Effort = f(Functional Size)

- Its equation takes this quantitative form:

$$\text{Effort} = a \times \text{Functional Size} + b$$

  - In practical terms, in this equation from the linear regression model (the straight blue line),
    - *a* represents the slope of the linear regression line
    - *b* represents the point at the origin (that is, when the independent variable is = 0)
  - Stated differently:
    - The slope <u>*a*</u> represents the increase of unit(s) of effort for an increase of 1 unit of functional size.

A Productivity Model: An ISO 15939 Measurement Information Model

- In terms of measurement units, this equation then corresponds to:

  - Effort (in hours) =
    - *a* (hours/Function Point) x Functional size (in Function Points)
    - + *b* (hours at the origin when the functional size = 0).

  - When the mathematical expression is worked out with its measurement units, then the end result is, indeed, in hours

    - Therefore: <u>both the left- and right-hand sides of the equation have the same measurement unit</u> = ‘hours’.

A Productivity Model: An ISO 15939 Measurement Information Model

B- The organizational reference context

- A specific productivity model built within an organization with the set of completed projects from this organization would then become the organizational reference context.
- Such a productivity model built from the organization's own past projects can then be used for estimating the next project for this organization. This productivity model will then:
  - provide a specific estimate that would be directly on the linear regression line, as well as
  - provide various elements of information on the quality characteristics of this model (such as its R^2, MMRE, etc.), which could be used as additional elements to make a decision on whether or not the selected estimate for the specific project would be:
    - above the regression line (i.e. more costly)
    - on the regression line
    - below the regression line (i.e. less costly).

A Productivity Model: An ISO 15939 Measurement Information Model

- With respect to the ISO 15939 Measurement Information Model,
  - the regression model corresponds to the ‘Analysis Model’,
  - the specific outcome of the regression model would be the ‘Indicator’, and
  - the set of information from the specific productivity model built by this organization would correspond to the ‘Interpretation’ context, while
  - the standard statistical technique of linear regression, which forms the basis for the *organizational* reference context would also be part of the Interpretation context.

Agenda

This chapter covers:

- The difference between a number and an analysis model.
- The Measurement Information Model in ISO 15939:
  - its metrology-related perspective
  - its analysis perspective for the quantification of relationships.
  - Examples of these differences within a Measurement Information Model.
- **A Metrology Design ≠ A Quantification Model of Relationships**
  - Example 1: the measurement of a single attribute
  - Examples 2: the quantification of relationships across attributes and entities.

© 2010 Alain Abran - Software Metrics & Software Metrology

A Metrology Design ≠ A Quantification Model of Relationships

- In ISO 9126: there are close to 80 attributes identified as required to be measured as necessary for the + 250 derived measures proposed to quantify the 3 ISO 9126 quality models, the 10 corresponding quality characteristics and the 27 quality sub-characteristics.

  - The measurement of one of these attributes, the ‘function’, is necessary for 38 different derived measures, while another one, the ‘user pauses’, is needed only in a single derived measure.

  - This section presents the outcomes of an exercise carried out in a graduate course where it was required to select an attribute from any of the 80 attributes in ISO 9126 and to design a corresponding measurement method.

Agenda

This chapter covers:

- The difference between a number and an analysis model.
- The Measurement Information Model in ISO 15939:
  - its metrology-related perspective
  - its analysis perspective for the quantification of relationships.
  - Examples of these differences within a Measurement Information Model.
- A Metrology Design ≠ A Quantification Model of Relationships
  - Example 1: the **measurement of a single attribute**
  - Examples 2: the quantification of relationships across attributes and entities.

A Metrology Design ≠ A Quantification Model of Relationships

- Using the measurement design methodology presented in chapter 2, and in particular the Design template presented in section 3, the graduate students came up with 2 very distinct types of design:
  - A design corresponding to the metrology related part of ISO 15939,
  - A design which, instead, took the perspective of the analysis of relationships and came up not with the design of a base measure, but with a quantification model of relationships across entities and attributes.

- An example of each type of designs are presented next, not because these 2 designs are complete and finalized, but only to illustrate that designers of software measures must beware that measurement and quantification are very distinct concepts, and have different properties.

Example 1: A Metrology Design of a Single Attribute

- **A- Example 1: Design of a base measure**
  - In ISO 9126 the 'number of cases' is necessary in 38 distinct derived measures. To obtain the 'number of cases' as a measurement result, it is necessary to have a well defined attribute of what is a 'case', and this definition should preferably be the same for each of the 38 distinct derived measure. This design has been described in more details in (Ozcan Top, 2009).

Step 1: Determination of the measurement objectives

  - The specific objective selected was the design of a measurement method for the size of a 'case'.
  - The results of the measurement method were intended to be used in the derived measures for the ISO 9126 characteristics and sub-characteristics listed in Next Slide.

Example 1: A Metrology Design of a Single Attribute

Characteristic	**Subcharacteristic**	**Sample Measurable Attributes**
Reliability	Recoverability	Availability
Functionality	Interoperability	Data exchangeability
Usability	Understandability	Demonstration Accessibility in use
	Learnability	Help frequency
	Operability	Customizability
Maintainability	Analyzability	Status monitoring capability
	Changeability	Parameterized modifiability
	Stability	Change success ratio
	Testability	Availability of built-in test function
Portability	Installability	Ease of installation

Table 2 Characteristics and Sub-characterictics requiring measurement of ‘cases’

Example 1: A Metrology Design of a Single Attribute

Step 2: Characterization of the concepts to be measured

- The characterization of the concept to be measured requires the definitions and the decomposition of such a concept.
- The characterization of a concept should initially be based on the findings from a literature review: 29 relevant references were identified, including:

- ISO FCD 24765 Systems and software engineering – Vocabulary.
- ISO 26514 Systems and Software Engineering - Requirements for Designers and Developers of User Documentation.
- ISO 19761:2003: Software Engineering - COSMIC-FFP: A Functional Size Measurement Method

Example 1: A Metrology Design of a Single Attribute

- **Definition and decomposition of the concept**
  - A number of concepts were identified from the literature review. Xt slide.
  - From these, the concept of 'action' was identified as the central one from a measurement perspective.

- **Definition of the Sub Concepts**
- 3 key sub concepts were identified:
  - **Input action:** "Any item, whether internal or external to the project that is required by a process before that process proceeds". "Data received from an external source"
  - **Output action:** "Data transmitted to an external destination". "A product, result, or service generated by a process."
  - **System action "**Set of interrelated or interacting activities which transforms inputs into outputs".

Example 1: A Metrology Design of a Single Attribute

Definitions from the literature review (Ozcan Top, 2009)

Case is defined by ISO 24765 as: "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".

Use Case is the description of the interaction between an Actor (the initiator of the interaction) and the system itself. It is represented as a sequence of simple steps. Each use case is a complete series of events, described from the point of view of the Actor.

Actor, Main scenario, Alternative Paths (Extensions), and Exceptions are the concepts that will be used as a basis of the measurement:

An Actor is "someone or something outside the system that either acts on the system – a primary actor – or is acted on by the system – a secondary actor. An actor may be a person, a device, another system or sub-system, or time. Actors represent the different roles that something outside has in its relationship with the system whose functional requirements are being specified."

Preconditions define all the conditions that must be true before the initiation of the use case.

Main scenario is the description of the main success scenario in a sequential order. **Action** is the element of a step that a user performs during a procedure.

Post-conditions "describe what the change in state of the system will be after the use case completes. Post-conditions are guaranteed to be true when the use case ends"**.**

Alternative paths; "Use cases may contain secondary paths or alternative scenarios, which are variations on the main theme. Each tested rule may lead to an alternative path and when there are many rules the permutation of paths increases rapidly.

Sometimes it is better to use conditional logic or activity diagrams to describe use case with many rules and conditions."

Exceptions, is the place "what happens when things go wrong at the system level are described, not using the alternative paths section but in a section of their own." An example of an alternative path would be: "The system recognizes cookie on user's machine", and "Go to step 4 (Main scenario)". An example of an exception path would be: "The system does not recognize user's logon information", and "Go to step 1 (Main path)"

Example 1: A Metrology Design of a Single Attribute

Step 3: Design of the Meta-Model

- Figure below presents the meta-model proposed to illustrate the relationships across the concepts and sub-concepts selected to characterize the size of 'cases'.

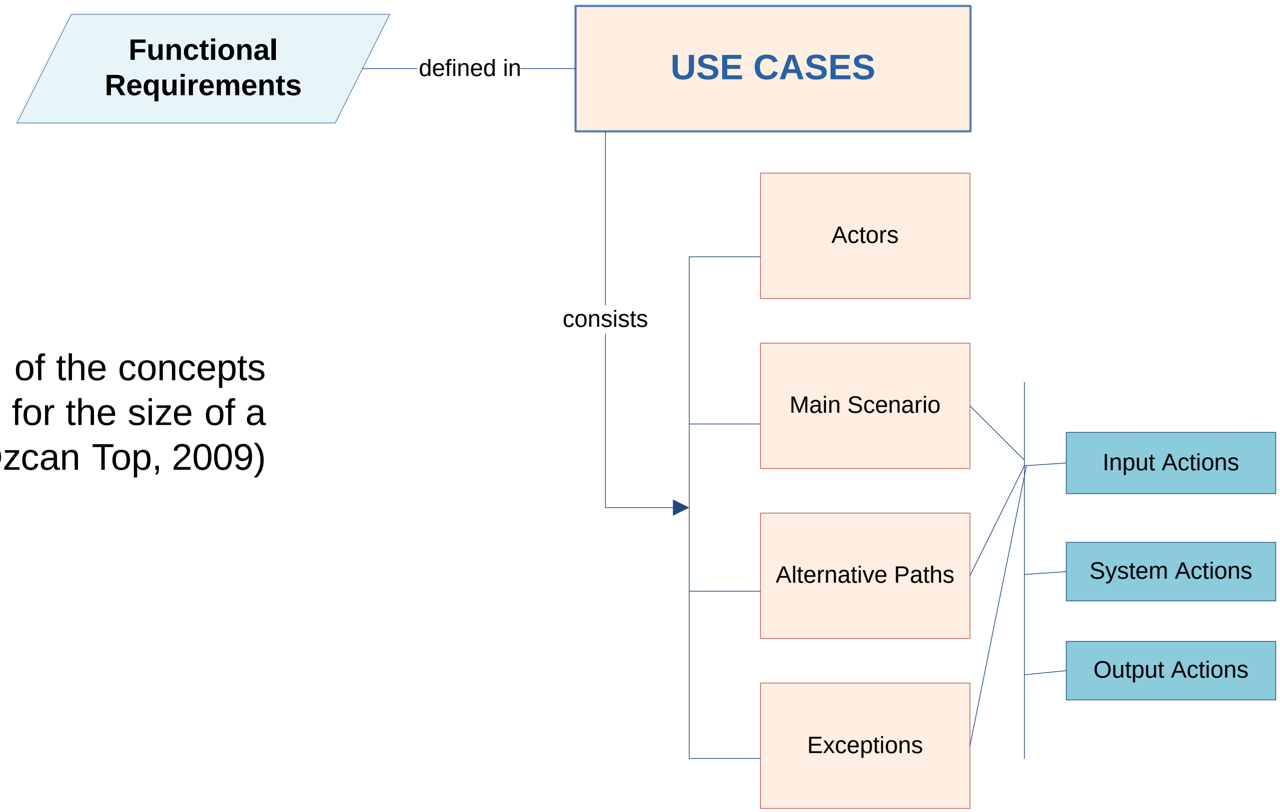


Figure 8: Meta-model of the concepts and sub-concepts for the size of a 'case' (Ozcan Top, 2009)

Example 1: A Metrology Design of a Single Attribute

Step 4: Assignment of Numerical Rules

- **The empirical description (and measurement unit)**
  - The size of a 'case' was defined as the addition of the Input Actions, System Actions and Output Actions.
  - According to this measurement function, each action type (Input Action, System Action, and Output Action) is assigned next a numerical size of 1 Action Unit (AU).
- **Mathematical Expression(s)**
  - The above empirical description can now be expressed as a mathematical expression:
  - Size of a Case = (Input Actions) + (System Actions) + (Output Actions)
- **Measurement Scale Type**
  - AU (Action Unit = 1) has a ratio scale type which means it can be used in statistical analysis and mathematical calculations.
  - These numerical assignment rules are presented in the next slide.

Example 1: A Metrology Design of a Single Attribute

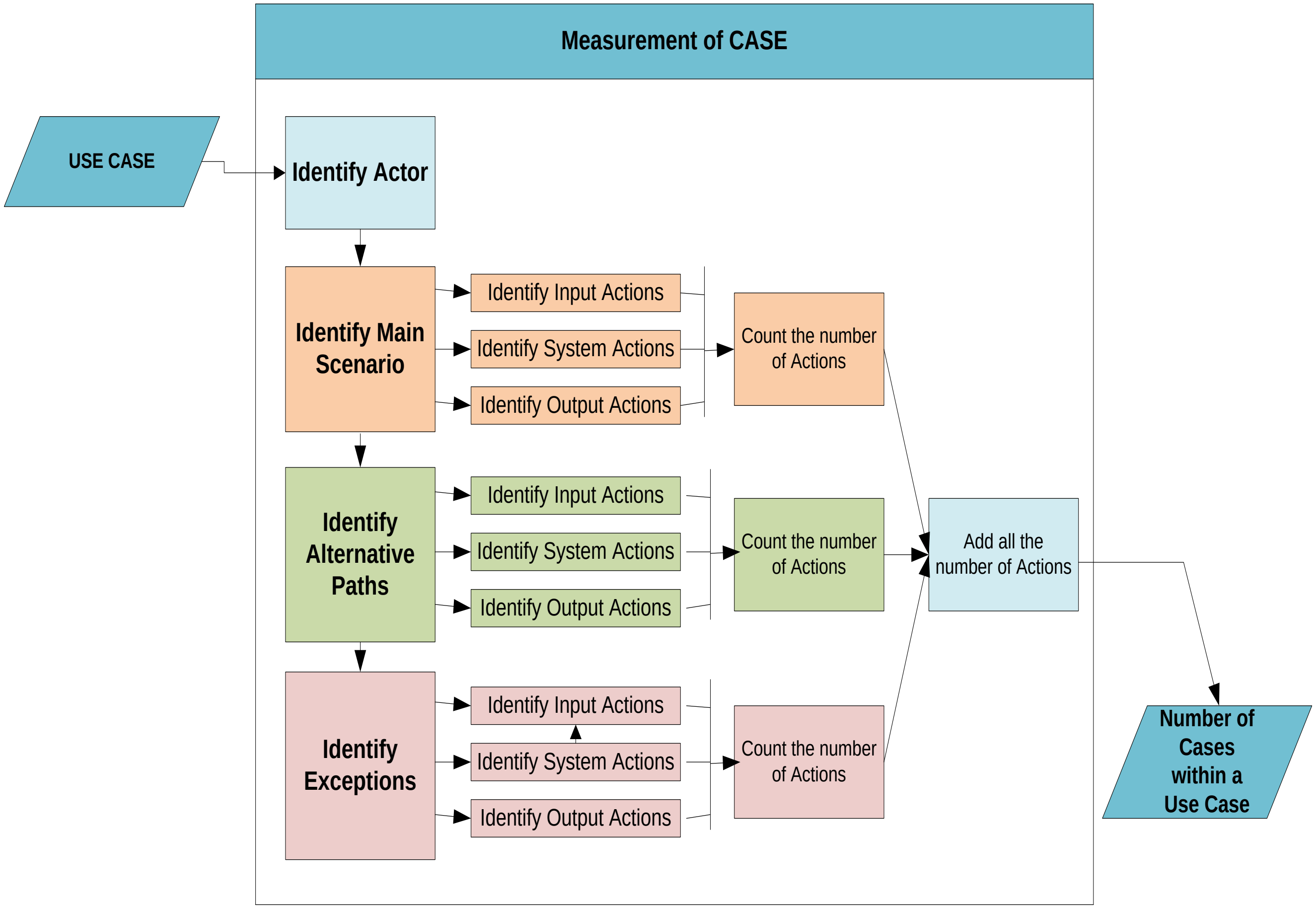


Figure 9: Measurement model of the size of cases within a Usecase (Ozcan Top, 2009)

Example 1: A Metrology Design of a Single Attribute

- Note: the example presented here has been designed by a single person, and should therefore be considered strictly as a first draft which should go through a number of iterations before reaching a certain level of maturity as a measurement method.

- Part 3 of the book presents in chapters 11 and 12 the full design process of a measurement method that has gained an international consensus as a software measurement method that is, ISO 19761: COSMIC.

Agenda

This chapter covers:

- The difference between a number and an analysis model.
- The Measurement Information Model in ISO 15939:
  - its metrology-related perspective
  - its analysis perspective for the quantification of relationships.
  - Examples of these differences within a Measurement Information Model.
- A Metrology Design ≠ A Quantification Model of Relationships
  - Example 1: the measurement of a single attribute
  - Example 2: the **quantification of relationships across attributes and entities.**

© 2010 Alain Abran - Software Metrics & Software Metrology

Examples: a Metrology Design & a Quantification Model of Relationships

- **B- Example 2: Design of a quantification model of relationships (across entities and attributes) (Dikici 2009)**
  - In ISO 9126, the number of 'error messages' is necessary to measure the 'Efficiency' and 'Resource Utilization'.

  Step 1: Determination of the measurement objectives
  - The specific objective selected was the design of a measurement method for the efficiency of 'error messages'.
  - The measurement results are intended to be used in the derived measures for the following ISO 9126 characteristics and sub-characteristics.
    - Usability
    - Operability
    - Understandability
    - Learnability

Examples: a Metrology Design & a Quantification Model of Relationships

- **Step 2: Characterization of the concepts to be measured Definition and Decomposition of the Concepts**
  - A number of concepts were identified from the literature review. Two of the main concepts identified were:
  - Message Appearance
  - Message Content.
  - In turn, each of these concepts can be decomposed in a number of sub-concepts – see Next Figure and Table.
- **Step 3: Design of the Meta-Model**
  - The identification of the relationships across the concepts and sub-concepts are illustrated in Next Figure.

Examples: a Metrology Design & a Quantification Model of Relationships

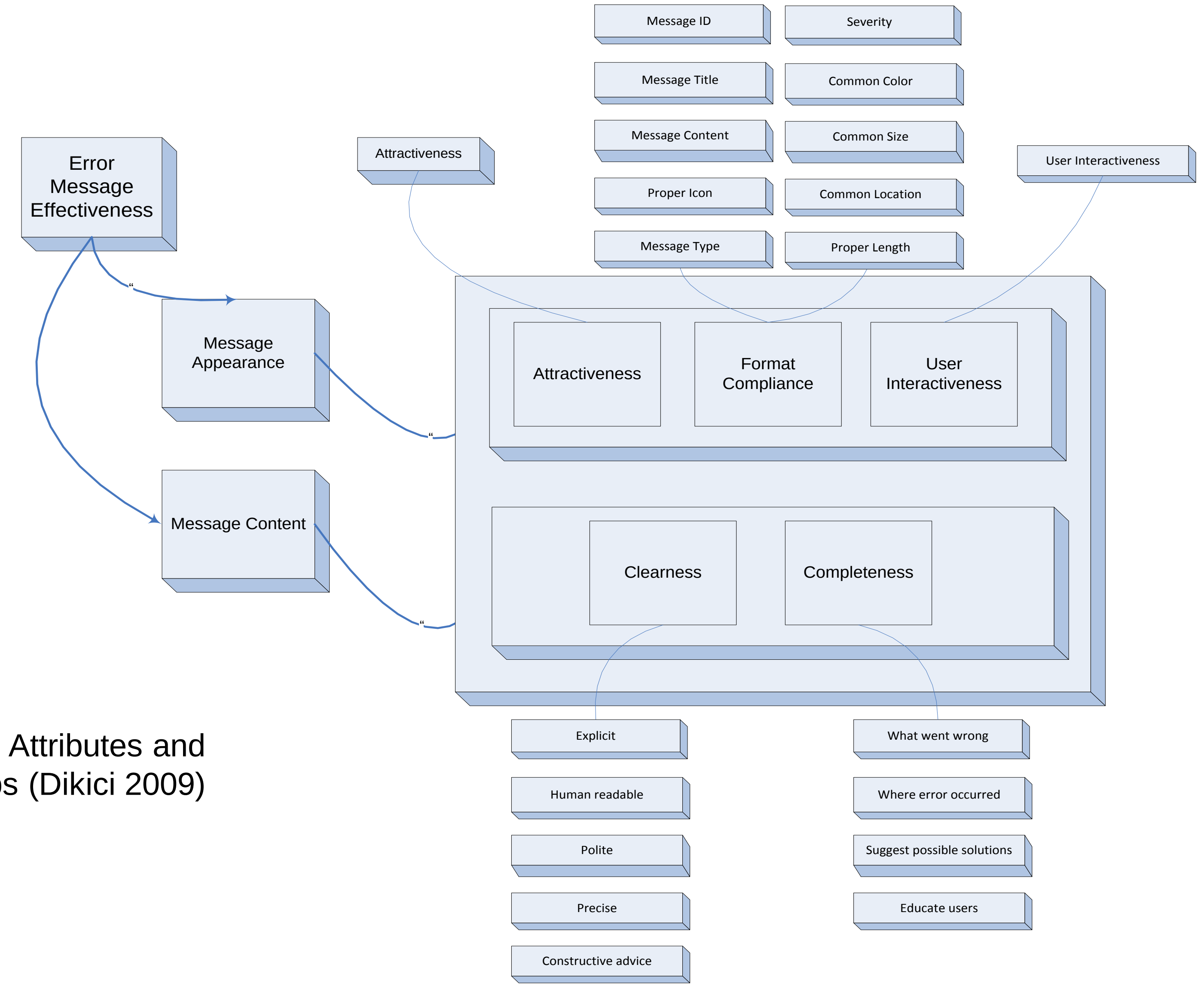


Figure : Attributes and Relationships (Dikici 2009)

Examples: a Metrology Design & a Quantification Model of Relationships

Concept	Sub-concept	Weight	Rank ranges
Message Appearance			
Attractiveness		10	
	Attractiveness		0-10
Format Compliance		40	
	Message ID		0/4
	Message Title		0/4
	Message Content		0/4
	Proper Icon		0/4
	Message Type		0/4
	Severity		0/4
	Common Color		0-4
	Common Size		0-4
	Common Location		0-4
	Proper Length		0-4
User Interactiveness		5	
	User Interactiveness		0-5
Message Content			
Clearness		25	
	Explicit		0-5
	Human-readable		0-5
	Polite		0-5
	Precise		0-5
	Constructive advice		0-5
Completeness		20	
	What went wrong		0-5
	Where error occurred		0-5
	Suggest possible solutions		0-5
	Educate users		0-5
		Total	
		Effectiveness %	

Table: The numerical assignment structure with weights and ranges of ranking (Dikici 2009)

Examples: a Metrology Design & a Quantification Model of Relationships

- **Empirical description**
  - The ‘effectiveness’ of an error message was defined as the quantification of both the appearance and the content or error message, on the basis of the quantification of each of their own sub-concepts, as illustrated in the meta-model presented in previous Figure.
  - All of these sub-concepts were themselves quantified individually using their own set of rankings assigned by the person in charge of evaluating the effectiveness of the error message.
  - For some of these sub-concepts the ranking selected were from o to 4, others from 0 to 5 and for some other ones, from 0 to 10.

Examples: a Metrology Design & a Quantification Model of Relationships

- **Mathematical Expression(s)**
  - The ‘effectiveness’ of an error message is calculated based on measuring the sub concepts presented in previous Figure.
  - The sub concepts are to be measured based on the rules specified in previous Table.
  - The 4 concepts are each assigned a relative weight (as a percentage).
  - The 21 sub concepts (within these 4 concepts) are next assigned a range of rankings, starting at 0, and up to 4, 5 and 10 – see previous Table.
    - In this specific numerical assignment rule, each sub-concept has an equivalent range within a concept (e.g. from 0-4 for the 10 sub-concepts participating to the upper concept which itself was assigned a weight of 40).

Examples: a Metrology Design & a Quantification Model of Relationships

- **Measurement Scale Type**
  - The numerical structure above is often used in practice in the evaluation of software quality based on a number of concepts and sub-concepts.
    - However, being used often in practice is no guarantee that this is the most appropriate mathematical structure.
  - In particular, the scale type of the end results of this set of numerical assignment rules is challenging to determine without ambiguity:
    - The intervals are in increments of 1 (from 0 to 4, for example), but there is no explicit definition of what is an interval of 1, and no explicit and rigorous definitions that subsequent intervals from 2 to 4 are indeed equal intervals.
    - Next, in practice, the selection of an interval is typically judgmentally based, and such a selection would often vary across people selecting a specific value, and may even vary if he same person was to select again a value let us say a week later.

Examples: a Metrology Design & a Quantification Model of Relationships

- In practice the corresponding values can certainly be considered as ordering values:
  - But considering them as being on a ratio scale type would be somewhat far stretched.
- It must be observed that each of the 21 sub-concepts in Table 3 is different, and if they were measured with an adequate design, they would each have their distinct measurement units and measurement scales: you would not then be able to add them up (since they do not have the same measurement units).

© 2010 Alain Abran - Software Metrics & Software Metrology

Examples: a Metrology Design & a Quantification Model of Relationships

- Therefore, adding up the values assigned to anyone of the 21 sub-concepts does not correspond to a measurement exercise, whether or not they have been multiplied by a ‘weight’ (or ‘points’ as in a number of software ‘metrics’).
  - Adding them up is a quantification, but without the rigor and meaningfulness of measurement with the rigor of metrology.
- This is typical of any quantification whereas weights and ‘points’ are somewhat arbitrarily assigned.
  - This will be illustrated in more details in chapter 8 in the analysis of the Function Points method and in chapter 9 in the analysis of the Usecase Points.

Summary

This chapter has presented:

- The difference between a number and an analysis model.
- The Measurement Information Model in ISO 15939:
  - its metrology-related perspective
  - its analysis perspective for the quantification of relationships.

  and various examples of these differences within a Measurement Information Model.
- Examples of the designs of:
  - The measurement of a single attribute – base measures
  - The combination of multiple single attributes – derived measures
  - The difference between a productivity ratio & a productivity model
  - The quantification of relationships across attributes and entities.

- Additionnal material

A Productivity Model: An ISO 15939 Measurement Information Model

- For the linear regression model (as a standard reference model), a number of the well known evaluation criteria of such statistical models are available in the literature, such as:

- *Coefficient of determination (R2)*
  - The coefficient of determination (R2) describes the percentage of variability explained by the predictive variable in the linear regression models.
    - This coefficient has a value between 0 and 1: an R2 close to 1 indicates that the variability in the response to the predictive variable can be explained by the model, i.e. there is a strong relationship between the independent and dependent variables.

A Productivity Model: An ISO 15939 Measurement Information Model

- *Error of an estimate – Error*
  - The effort of an estimate (i.e. Error = Actual – Estimate) represents the error of the estimation model on a single project. For example, the difference between the known effort of a project completed (i.e. Actual) versus the value calculated by the model (i.e. Estimate).

- *Relative Error (RE)*
  - The relative error (*RE*) corresponds to the Error divided by the Actual.

- Magnitude of relative error: *MRE:*

$$\underline{MRE} = |RE| = \left|\frac{Actual - Estimate}{Actual}\right|$$

- Mean magnitude of relative error: *MMRE:*

$$\underline{MMRE} = \overline{MRE} = \frac{1}{n}\sum_{i=1}^{n} MRE_i$$

A Productivity Model: An ISO 15939 Measurement Information Model

- <u>The root of the mean square error: *RMS:*</u>

$$\underline{RMS} = \left(\overline{SE}\right)^{1/2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Actual_i - Estimate_i\right)^2}$$

- <u>The relative root of the mean square error: *RRMS:*</u>

$$\underline{RRMS} = \overline{RMS} = \frac{RMS}{\frac{1}{n}\sum_{i=1}^{n} Actual_i}$$

- <u>*Predictive quality of the model*</u>
  - The prediction level of an estimation model is: $PRED(l) = \frac{k}{n}$, where *k* is the number of projects in a specific sample of size n for which $\mathrm{MRE} \leq l$.

A Productivity Model: An ISO 15939 Measurement Information Model

- **<u>Some interpretations of the values of these evaluation criteria</u>**
  - The smaller the *RMS* or RRMS, the better the prediction level.
  - In the software engineering literature, an estimation model is generally considered good when:
    - the MRE (Mean Relative Error) is within +/-25% for 75% of the observations, or
    - PRED(0.25) = 0.75.

Examples of different perspectives in ISO 15939

- In practice, there is no guarantee that what can be measured adequately at the level of base and derived quantities does indeed represent the concepts and relationships that the analysis part of the Measurement Information Model attempts to quantify.
  - An example of this is the *maintainability* characteristic in ISO 9126, which is:
    - not strictly limited to the software entity itself, but is
    - implicitly related to the entity ‘effort required to maintain such software at a later time’.

- Chapter 10 will present an example with the ISO 9126 quality models whereas such a link is not yet mature and where much work remain to clarify the links between its measurement part and the quantitative analysis.

A Productivity Model: /

Table 1: Recording rules for the 'Effort' variable – Source: www.isbsg.org

TIME RECORDING METHODS	WORK EFFORT BREAKDOWN
Method-A: Staff Hours (Recorded) The daily recording of all of the WORK EFFORT expended by each person on Project related tasks. As an example, where a person who works on a specific project from 8am until 5pm with a 1 hour lunch break will record 8 hours of WORK EFFORT. **Method-B:** Staff Hours (Derived) It is possible to derive the WORK EFFORT where it has not been collected on a daily basis as in Method-A. It may have only been recorded in weeks, months or years. **Method-C:** "Productive" Time Only (Recorded) The daily recording of only the "productive" effort, (including overtime), expended by a person on project related tasks. Using the same example as used in Method-A above, when the "non-productive" tasks have been removed, (coffee, liase with other teams, administration, read magazine, etc.), only 5.5 hours may be recorded.	Data collected about the people whose time is included in the project work effort. **Level 1: Development Team** Those responsible for the delivery of the application under development. The team or organization, which specifies, designs and/or builds the software. It typically also performs testing and implementation activities. It comprises: Project Team Project Management Project Administration Any member of IT Operations specifically allocated to the project **Level 2: Development Team Support/IT Operations** Those who operate the IT systems that support the end-users and are responsible for providing specialist services to the Development Team, (but not allocated to that team). Support comprises: Data Base Administration Data Administration Quality Assurance Data Security Standards Support Audit & Control Technical Support Software Support Hardware Support Information Centre Support **Level 3: Customers / End Users** Those responsible for defining the requirements of the applications and sponsoring/championing the development of the application. Also the software's end users. The relationship between the project customer and the software's end users can vary, as can their involvement in a software project. It comprises: Application Clients Application Users User Liaison User Training