XVIII IMEKO WORLD CONGRESS
Metrology for a Sustainable Development
September, 17 – 22, 2006, Rio de Janeiro, Brazil

# INTRODUCTION TO "SOFT" METROLOGY

*Jean-Claude Krynicki*

European Quality and Metrology Manager
Agilent Technologies Service and Support Unit, Palaiseau, France, jean-claude_krynicki@agilent.com

**Abstract:**
Outside of the classical physical and chemical metrologies, there is a large field of measurement, which is called dimensionless or "soft " metrology by analogy with hard and soft sciences.
The objective of this presentation is to provide an overview of this field, to analyze the methods used and their linkage to classical metrology.
The author will take 3 examples, an historical one, the measurement of human character, and two modern research fields, software usability and customer satisfaction.
He will also show the recent efforts from the European Commission, as part of the New Emerging Sciences and Technologies program (N.E.S.T.), to increase the objectivity of these measurements.

**Keywords:** soft metrology, usability, customer satisfaction.

## 1. INTRODUCTION

The new version of ISO 9001-2000 is considering the measurement of customer satisfaction as an integral part of metrology. The logic behind this requirement is coming from the situation of the 70% certified companies operating within the service field. In most case they have not measuring instruments at all. The only measurement they should make to validate their operation is the evaluation of customer satisfaction.

With this example, dimensionless or "soft" metrology has made its official introduction in the field of metrology as a discipline in itself and as an auditable requirement of paragraph 7.6 (Control of monitoring and measurement devices) [1]. In fact, since 1994, this requirement was part of auditor guides but the control process was not enforced [2].

Based on the success and universality of metrology there is a trend to apply measurement to any field of activity. This helps to provide an objective evaluation, to monitor trend or to perform benchmarks.

## 2. MEASUREMENT SCOPE

Soft metrology covers a broad range of mesurande, addressing all the aspects of measurement, outside of the traditional field of physical and chemical metrology :

- psychometric measurement or perceived feeling (color, taste, odor, touch),
- qualitative measurements (perceived quality, satisfaction, comfort, usability),
- econometrics and market research (image, stock exchange notation), sociometry (audience and opinion),
- measurements related to the human sciences: biometrics, typology, behavior and intelligence.

Measurement associated with perception were studied by comparison with associated physical results. In other words the human being is considered as a transducer and the analysis is focusing on creating a scale, selecting the panel of sensation experts, defining sensitivity, repeatability and reproducibility and comparing the results with the one obtained by methods from traditional metrology. Therefore this field of study was promoted in metrology congresses, either as an opening to new applications [3] or as a specific section [4] [5][6]. In a same way "soft" metrology departments were created at NIST and NPL.

Our proposal is to investigate other fields of "soft" metrology activities where everything needs to be created from the definition of the mesurande to the design and validation of the measurement model. The author will present 3 fields of activity he has experimented, either within professional practice or personal interest. The results are debated but they allow to create a general view of the metrology issues associated with these fields of investigation.

## 3. EXAMPLES

### *3.1. The measurement of human character*

Since the beginning of antiquity there is a belief that each human being has his behavior influenced by a psychological pattern called the character. This pattern can be classified into specific categories; each category can be evaluated within a scale. The character is supposed to be stable during the life of the subject. The subjects

having the same character will react the same way in front of the same situations.

Among 15 ways to analyze the character, about 5 can be considered as scientific, based on factor analysis using large populations such as the systems of Catell, Heymans or the Big Five. Despite one century of researches there are still divergence in the items composing such models.
According to Catell [7] the knowledge of character allows to define a set of specific parameters in the equation:

Pij = sa F ai + sb F bi + Sc F ci ….. + sj F ji

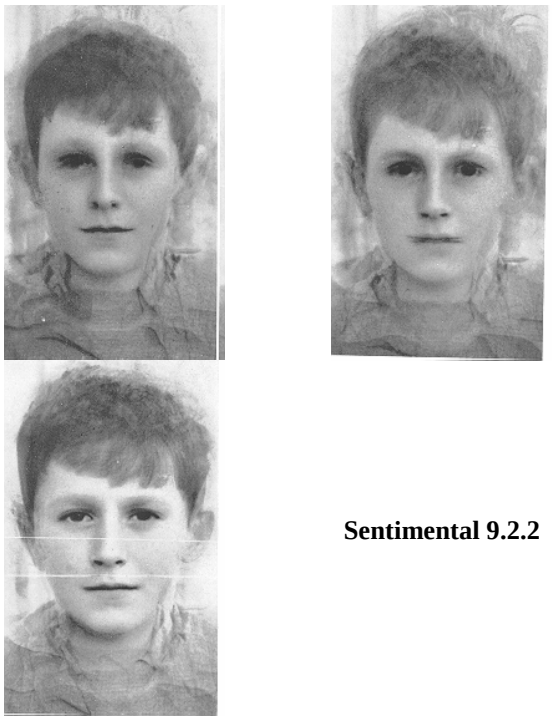
A subject (i) within a situation-stimulus (j) corresponding to the behavior under study (P) is activating several factors of character (Fa, Fb, Fc ….. + a residual Fj specific to the situation) with variable rates (sa, sb, sc … sj) which can be evaluated .
There are several "situation index" correlated to these factors. Statistical tests are also allowing to define a "sincerity index".
Several other models were proposed to identify the respective part of nature and culture within the behavior. As psychology is also taking care of abnormal behavior, some models are dealing with the balance between factors or their integration (Boven, Szondi)[7]. The general tendency is to reduce the number of factors to a small set allowing the quick identification of few representative profiles. However the diversity of characters, seen within the real population, lead to the addition of complementary factors.

**Example: the character analysis of Heymans – Wiesma – Le Senne**
We have selected this approach due to the size of the experimental population, the 70 years of researches, the several hundred publications and applications in the field of education and work.
The initial study was concerning the transmission of psychological profile from parents to children. After a selection of character traits from the usual vocabulary, a prototype questionnaire was created, based on 90 items, and applied to a population of 4000 subjects; the first results were compared with the evaluation of physicians knowing well the subjects. At the end of the study, 2500 robust cases were kept for analysis
Out of the study, 3 fundamental factors were identified, regardless of the intellectual level and socio-economic situation.

- Emotivity (E, nE) corresponding to the ability to react emotionally
- Activity (A, nA), linked to the ability to sustain an effort
- Repercussion or primarity - secondarity (P, S) , associated to immediate and short reaction in front of impressions (primarity) or delayed and prolonged impression (secondarity).

The distribution of the results shows 8 modes: Nervous (EnAP, Sentimental (EnAS), Bloody (nEAP), Phlegmatic (nEAS) Passionate (EAS), Choleric (EAP), Amorphous (nE,nA,P), Apathetic (nE,nA,S).

After this codification, several authors have tried to scale the contribution of each factor by refining the questionnaire. From that point, it was possible to create a profile associated, for example, with job soft skills, to be used in school orientation [8].
Several contributors have tried to complete the original models with their own factors to reflect the numerous psychological profiles not really represented by the initial model. A consensus was obtained with 12 factors.
Since Galton, there was also a trend in this area to define a morpho-type. Roger Mucchielli has applied the questionnaire to a population of 1000 teenagers, and has obtained the morpho-type by superposition of 16 pictures of each identical profile. Therefore this methods is not only able to "measure" the character, but also provide a composite visual picture of the "standards".



**Sentimental 9.2.2**

**Fig. 1 By superposing 16 pictures of subjects having the same result at the profile questionnaire and by focusing on the eyes, the nose and the mouth. It is possible to obtain a clean picture representing the profile (technique created by David Katz from the university of Stockholm).**

To conclude with this topic we can say that the mesurande was well identified. Le Senne has provided a clear definition of the character ("permanent skeleton of dispositions, building the mental structure of man") and other associated terms.
The procedure used to define the factors can be challenged; the selection of words used in the vocabulary to describe traits of character gives different results depending on the authors (5000 according to Allport, 650 according to Sheldon); therefore the factor analysis lead to different conclusions. Catell has found 16 factors (16PF test), Heymans 3, Myers and Briggs: 4 (MBTI test), Big Five test : 5. When the questionnaire is stabilized, the codification / measurement is biased because the

questions are only addressing the items of the model. Alfred Binet who created the first test of intelligence in 1905 has well described this tautology: “What is intelligence? This is what my test is measuring!”.
Units and scale are usually raw approximates; there is no evaluation of scale sensitivity. The only uncertainty provided is concerning the sample size. The main method, self-administered questionnaire, was not really validated or crosschecked with other techniques. The general trend is to capitalize on affirmation of recognized authors and to add a few variation on the model based on case studies.
However the results obtained by these codifications (and in some cases measurements) are enough for the level of expectation.

### *2.2. The measurement of software usability*

The interface between man and computer is an increasing concern. Some recent researches have tried to identify the ideal functionalities of such interface. After the release of the ISO 9241 standard, further university researches have created a questionnaire and a measurement scale [9] [10]. This is a good example, within a short period of time, on how to convert a vague concept into a precise definition, a set of attributes and then to create a measurement system.

Software usability was a debated concept since the introduction of German Ergonomic Dialog Design norm DIN 66234 Part 8, 1988. Most recent studies are defining usability in terms of:

- Effectiveness (the accuracy and completeness with which users achieve specified goals),
- Efficiency (the resources expended in relation to the accuracy and completeness with which the users achieve goals),
- Satisfaction (the comfort and acceptability of use),

Further breakdown of the concept into measurable parameters ended with 7 categories:

- **Suitability for the task**
  How well is the application adapted to the current activity; additional task are reduced to a minimum; the terminology used corresponds to the field of activity; the way data are entered follows the work process; the important commands to perform are easy to find; data presentation and reports are adapted to the task….

- **Self descriptiveness**
  The software messages are self-explanatory (feedback message, inputs confirmation messages, warning and error messages); explanations about field entries are retrievable; explanations are context sensitive.

- **Controllability**
  Navigation within the application is possible; switching between different menu levels is possible; the main menu is accessible from all screens; the software allows for task interrupt; keyboard shortcuts are possible (letter or command code).

- **Conformity with user expectations**
  Time spent by the software to perform an operation is predictable; the labels are consistent within the different modules; the same function keys are used for the same operations; next screen to appear in a given sequence can be predicted.

- **Error tolerance**
  When a mistake is entered, previous data entries are not lost; previous state can be restored; deleting data action triggers a confirmation message; data are checked on the fly; no system hangs are experimented; information to recover from errors is provided; error message are helpful to identify what went wrong.

- **Suitability for individualization**
  Individual preferences can be managed (forms, character size, menus, amount of information); speed of input devices and software response speed can be adjusted

- **Suitability for learning**
  The on line explanations are suitable for different level of knowledge, skills and needs for explanations

After some practice of application software customers using classical interviews it appears that the model proposed in the standard is very pertinent and does not miss any item mentioned by users. This measurement allows to structure the evaluation, to identify areas of improvement and to track effect of changes (relative measurement). The most challenging aspect is the evaluation of the “comfort” component, which is also an important research area in “soft” metrology. The user opinion is a biased with the expectations that a software application can remove any repetitive task and can be redesigned on the fly when expectations are changing. At the opposite, software development relies on initial specifications, language and architecture options selected out of these specifications, with little flexibility. Therefore this measurement technique is more suited for pilot evaluation and reduction of subjectivity in customer – supplier disputes (can be assessed by an independent expert).

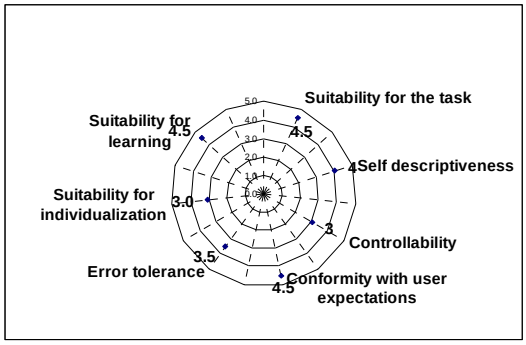


**Fig 2 An example of software usability measurement according to the IsoMetric process**

### *2.3. The measurement of customer satisfaction*

During a decade, several marketing research have tried to standardize a model of customer satisfaction based on a questionnaire and several statistical clusters. The first model, called the American Customer Satisfaction Index (ECSI-1994), was published by the University of Michigan and was composed of 6 elements (customer expectation, perceived quality, perceived value, customer satisfaction, customer complaint and customer loyalty). These statistical clusters were calculated using a partial least square (PLS) method. Later on, the European Commission has funded two research programs in this area, the European Customer Satisfaction Index (ECSI – 2000) and the European Satisfaction Index System (ESIS-2004); the model was refined by a split between product and service and the addition of Company Image.[11] .[12]

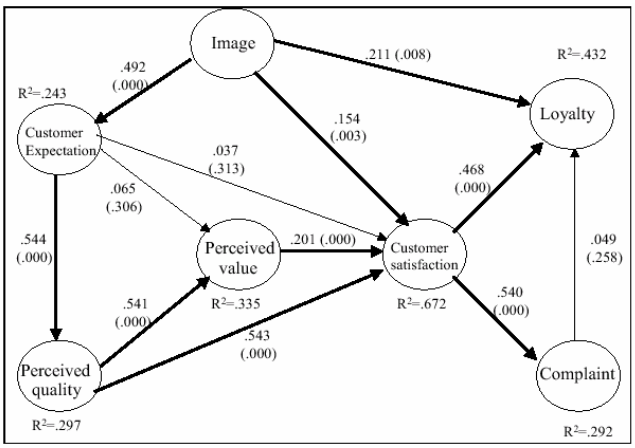



**Fig 3 An example of customer satisfaction index for the mobile phone industry**

The model is composed of several statistical clusters with a central indicator, the customer satisfaction (called the consumer confidence index in the US national survey), which provides a single value result. This approach allows to measure trends as well as to evaluate what should be improved.

The evaluation is performed using a survey questionnaire with several questions addressing each item of the model :

- Perceived product quality: Hardware product quality perceived by customer
- Perceived service quality: Software and human ware quality perceived by customer
- Company Image: Refers to the brand name and the kind of associations customers get from the brand/product/company.
- Expectation: Refers to the customer expectations before getting the product/service
- Perceived value: Perceived level of product quality relative to the price paid or the "value for money"
- Customer loyalty: Repurchase intention, price tolerance and intention to recommend products and services to others
- Complaints management: Refers to the intensity of complaints and the manner in which the company manages these complaints

In fact the real objective is more to evaluate the customer loyalty (re-purchasing potential) than satisfaction. Due to the large variations in the coefficient linking the items, the model is evaluated for a specific field of activity (automotive, airline, bank….).

The partial least square technique is generally adopted due to its stability. There are several variations in the questionnaire but the large population used, allows to identify any bias in the interpretation.

With this system made available by research, it is possible to evaluate customer satisfaction for a specific company as well as to compare the results to an aggregate of competitors.

Such research program was triggered and funded by national or European authorities to improve the reliability of satisfaction measurement. However, in practice, many companies are still creating their own questionnaire. The measurement uncertainty is usually limited to the sample contribution [13][14]. .

## 4. RECENT EVOLUTIONS

There is some reluctance, within the metrology community, to consider "soft" metrology as a field to be integrated within current practices.

The main argument is linked to subjectivity: using a questionnaire and getting a self-evaluation from the subject is introducing a bias difficult to evaluate. Other arguments are referring to cultural bias.

The ideal would be to perform physical measurement using sensors applied to a subject placed in a test situation.

The European Commission, as part of the New Emerging Sciences and Technologies program (N.E.S.T.), has funded a research program named "Measuring the Impossible"; the objective is to expand the metrology field using scientific methods within interdisciplinary research programs.

Recognizing the importance of physical and chemical measurements, the sponsors of this program are also explaining that "Science, business and government now present challenges to measurement which are intrinsically more complex, problematic and subject to interpretation". They mention that:

- "Many phenomena of significant interest to contemporary science are intrinsically multidimensional and multi-disciplinary, with strong cross-over between physical biological and social sciences.
- Products and services appeal to consumers according to parameters of quality, beauty, comfort, etc., which are mediated by human perception.
- Public authorities, and quasi public bodies such as hospitals, provide citizens with support and services whose performance is measured according to parameters of life quality, security or wellbeing".[15]

This is a clear request to expand the scope of metrology beyond physical and chemical measurements.

After the first run of selection of research proposals, we can mention some observations:

- This is a large and multi-disciplinary field of activity
- Very few European metrology institutes have submitted research proposals. Most of the proposals are coming from universities or specialized research centers (human sciences, economy, medicine, marketing etc….)
- The measurement systems proposed are very often composed of brain sensors (MRI, EEGs) combined with neural networks able to identify signal patterns. Uncertainty calculation for such systems is quite a challenge far beyond the issues metrologists are currently meeting with chemical components identification, vector signal analyzers or 3D coordinate machines.

However we are at a starting point and initiatives like the European Research Area Network for Metrology (MERA) are showing the way. Normative guidelines have also started to clarify measurement requirements.[16]

## 5. CONCLUSION

"Soft" metrology is a huge field of investigation. Metrologists can benefit from these researches. There is a possible cross-fertilization between the disciplines For example, the most common statistical tool used to measure customer satisfaction is the partial least square method (PLS) which was invented to improve chemical measurement.

Perception of colors, odors and touch are entry point in this field because human perception can be backed up by physical measurements.

Soft metrology sections were already created at NIST (USA) and NPL (UK).

## REFERENCES


[1] ISO 9001-2000 – Quality management systems – Requirements.

[2] A.F.A.Q. Auditor guide, 1994.

[3] J.F. Bassereau, R. Duchamp - Psychometric measurement of material hardness or how the measure physically the human perception of material hardness - 8th International Congress of Metrology - France- 1997

[4]Measuring associating physical instruments and new methods of sensory metrology of the touch, progress report of an interdisciplinary project – Mr Bassereau, Mr Deveau, Mr Duchamp 12th International Metrology Congress – Sensorial and immaterial metrology - Lyon, 2005.

[5] Metrology of wood carpentry visual aspect – Mrs Garay, Mme Lafon - 12th International Metrology Congress - Sensorial and immaterial metrology - Lyon, 2005.

[6] Perceived quality on large, deformable plastic body panels – Mrs Petri, Mr Boulakhrif - 12th International Metrology Congress - Sensorial and immaterial metrology - Lyon, 2005.

[7] Muchielli R., La caractérologie à l'ère scientifique (Character analysis at scientific age) - Edition du Griffon – 1961

[8] Caractérologie et integration sociale (Character analysis and social integration) in La caractérologie, volume 7 – Presses universitaires de France – 1965

[9] ISO 9241 – 1997, Ergonomic requirements for office work with visual display terminals, Part 11: Guidance on usability specification and measure

[10] Gunter Gediga, Kai-Christoph Hamborg, Ivo Duntsch – The IsoMetrics Usability Inventory, An operationalisation of ISO 9241-10 supporting summative and formative evaluation of software systems – 2004

[11] Michel Tenenhaus, Vincenzo Esposito Vinzi, Yves-Marie Chatelin, Carlo Lauro – PLS Path Modeling – 2002

[12] Michel Tenenhaus, Vincenzo Esposito Vinzi, Yves-Marie Chatelin, – ESIS, State of Art on PLS Path Modeling through the available software – 2002

[13] An example of "soft" metrology : the measurement of customer satisfaction - Jean-Claude Krynicki - NLA congress, Johannesburg - 2004.

[14] Jean-Claude Krynicki, Christophe Perruchet - Evaluation of uncertainties for customer satisfaction measurements - NCSLI 2006.

[15] N.E.S.T. 6th Framework Program, anticipating scientific and technological needs, Reference document on "Measuring the impossible" – European Commission December 2004

[16] AFNOR FDX 07-050 Metrology of immateriality — Measurement of expectation and perception — Recommendations for the control of expectation and perception measurements and their reliability – AFNOR 2003.