

Methods for Measuring Usability

Nigel Bevan and Ian Curson

National Physical Laboratory, Usability Services
Teddington, Middlesex, TW11 0LW, UK
Nigel.Bevan@npl.co.uk, Ian.Curson@npl.co.uk

ABSTRACT The tutorial explains the benefits of measuring usability as part of a user-centred design process, and introduces the participants to the methodology for usability measurement developed by the collaborative European ESPRIT MUSiC (Measurement of Usability in Context) project. The tutorial includes demonstration of the use of MUSiC tools, and class exercises to apply the methods to case studies.

KEY WORDS: Usability evaluation, metrics, usability engineering, video analysis

OBJECTIVES

In many organisations usability is ignored because there are no objective criteria for usability when developing and procuring products. The tutorial explains how usability can be operationalised as the prime quality goal: to provide quality in use (Bevan, 1995, 1996). Participants learn how to choose and apply appropriate cost-effective methods and tools in commercial projects at different stages of development, and how the methods relate to business goals. They gain experience in the use of tools for usability specification and measurement developed by the European MUSiC project (Bevan and Macleod, 1994), implementing the principles of ISO 9241-11.

The tutorial is intended for anyone who wishes to specify or measure usability as a quality objective during product development or to evaluate usability when purchasing. The emphasis is on explaining when it is beneficial to use evaluation procedures which provide usability measures in addition to design feedback.

Usability measurement provides benefits not available from qualitative evaluation methods. Measurement can be used to:

- predict, ensure and improve product quality
- control and improve the production processes
- decide on the acceptance of a software product
- select a product from among alternative products

CONTENT

The tutorial starts by introducing the benefits of setting and testing usability goals in design, and shows how usability goals can be derived by decomposing overall business goals for the product. In particular circumstances, the goals for usage in a particular context may be in terms of user performance, attitude or cognitive workload. The following topics are covered:

1. The benefits of usability measurement are related to evaluation procedures in ergonomics, software quality and quality management. Different approaches to usability are reviewed, and emphasis is given to the benefits of defining usability in terms of meeting the needs of users by providing quality of use.

2. The importance of specifying the context of use is explained, and using the Usability Context Analysis Guide (Macleod, 1994) the class carries out an exercise to specify the intended context of usage, the actual context of measurement and a measurement plan including specific usability goals.

3. Use of performance measures: including a description of the MUSiC Performance Measurement Method (Macleod et al 1997) and DRUM (Macleod and Rengger, 1993) which can be used to derive user performance metrics and diagnostic information from an analysis of a videotape of interaction.

4. Use of attitude questionnaires: including a description of the validated SUMI scale which can be used to measure six components of attitude towards the interaction and identify problem areas of the interface.

5. Measurement of cognitive workload using questionnaires is explained.

6. Case studies: use of each of the tools is illustrated by case studies and class exercises which draw from the experience of the use of the methods in a wide variety of commercial applications.

Performance Measures (DRUM Tool)

The MUSiC User Performance Method is a technique for measuring the usability of a product by measuring the extent to which specific users of the product achieve specific goals in a specific environment.

The data on which the usability metrics are based is recorded from a variety of sources during evaluation sessions, in which subjects representative of the user community of the product being tested, carry out tasks that are typical of those the product's users undertake, in an environment similar to that in which the product is normally used. The typical characteristics of the product's users, the tasks they do, and the environment in which they do them are identified by following the guidance provided in the Usability Context Analysis Guide.

The analytical load on the analyst is reduced by firstly providing information, in a document called A Guide to Analysing Usability Sessions, that describes how an analyst can objectively identify particular user actions during the sessions, and secondly by providing a software tool, the Diagnostic Recorder for Usability Measurements, DRUM, that helps the analyst build a time-sequence log of the pre-defined user actions. From these logs of the sessions the DRUM can automatically calculate usability metrics for individual users and group the results of similar users together.

Attitude Measures: SUMI

The assessment of a computer system's usability should involve measuring not only aspects of users' performance, but also how users feel about the system they are using. The Software Usability Measurement Inventory (SUMI) has been designed specifically for this purpose. SUMI is designed to investigate users' perceptions of the quality of software systems. It is applicable to many sorts of applications. SUMI gives the investigator:

- A Global Usability measure
- Five Usability subscale measures
- A high level problem diagnosis.

SUMI offers the investigator norms based on large samples, and the possibility of comparing the system being tested either against generic usability profiles, or against the usability profile of another system.

Cognitive Workload Measures

Efficient use of the computer is intrinsic to usability. This means that in assessing the usability of software it is necessary to investigate both user performance and the effort a user invests in using the software. If a good performance can only be achieved at the expense of a high amount of invested effort, a system is not usable. MUSiC has objective and subjective measurement techniques suited for use in a wide variety of applications. Participants are introduced to the different measures of cognitive workload, and the situations in which the different measures can be used together or separately.

REFERENCES

- Bevan N (1995) Usability is quality of use. In: Anzai & Ogawa (eds) Proc. 6th International Conference on Human Computer Interaction, July 1995. Elsevier.
- Bevan N (1995) Measuring usability as quality of use. *Journal of Software Quality*, 4, 115-130.
- Bevan N, Macleod M (1994) Usability measurement in context. *Behaviour and Information Technology*, 13, 132-145.
- Macleod, M (1994) Usability in Context: Improving Quality of Use. In: G Bradley and HW Hendricks (eds.) *Human Factors in Organizational Design and Management - IV - Proceedings of the International Ergonomics Association 4th International Symposium on Human Factors in Organizational Design and Management*, (Stockholm, Sweden, May 29 - June 1 1994). Amsterdam, Elsevier.
- Macleod M, Bowden R, Bevan N and Curson I. (1997) The MUSiC Performance Measurement Method. *Behaviour and Information Technol.*, in publication
- Macleod M and Rengger R. (1993) The Development of DRUM: A Software Tool for Video-assisted Usability Evaluation. In: *People and Computers VIII, Proceedings of the HCI'93 Conference* (Loughborough, UK, Sept 1993), CUP.