Proceedings of HCI International 2009, San Diego, California, USA.

Extending Quality in Use to Provide a Framework for Usability Measurement

Nigel Bevan

Professional Usability Services, 12 King Edwards Gardens, London W3 9RG, UK
mail@nigelbevan.com

ISO has recently developed a new more comprehensive definition of quality in use, which has usability, flexibility and safety as subcharacteristics that can be quantified from the perspectives of different stakeholders, including users, managers and maintainers. While this provides a more complete set of requirements for operational use of a product, it also presents new challenges for measurement.

Keywords: standards, usability, quality in use, requirements, measurement, safety

1 Usability and quality in use

A traditional view of usability that is popular among product developers is that it is the attributes of the user interface that makes a product easy to use. This is consistent with one of the views of usability in HCI, for example in Nielsen's [22] 1993 breakdown where a product can be usable, even if it has no utility:

- system acceptability
  - social acceptability
  - practical acceptability
    - cost
    - compatibility
    - reliability
    - usefulness
      - utility
      - usability

This view of usability is also consistent with the first ISO definition of usability as part of software quality in ISO/IEC 9126 (1991):

> **Usability**: a set of attributes that bear on the effort needed for use, and on the individual assessment of such use, by a stated or implied set of users.

This definition of user interface usability contrasts with the system perspective of usability defined from an ergonomic point of view in ISO 9241-11 (1998):

Usability: the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

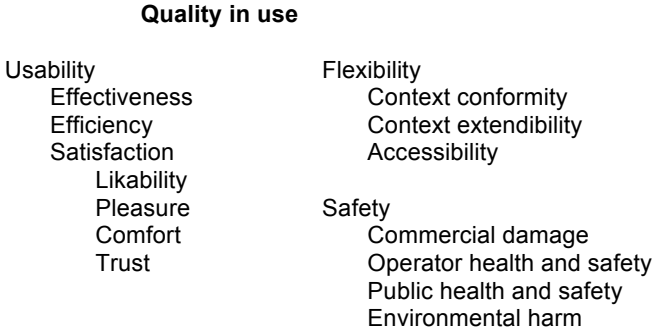
This wider interpretation of usability was incorporated in the revision of ISO 9126-1 (2001), renamed "quality in use" as it is the user's perspective of the quality when using a product [3]. The software quality characteristics: functionality, reliability, efficiency, usability, maintainability and portability contribute to this quality.

2 The new definition of quality in use

In 2006 the US standard for a Common Industry Format for Usability Test Reports (CIF) [1] was adopted by ISO as part of the revised Software product Quality Requirements and Evaluation (SQuaRE) set of standards [18]. As potential users of the CIF had originally expressed a preference for the term "usability" rather than "quality in use", the ISO 9241-11 definition of usability was retained when the CIF became part of this series.

When the ISO/IEC 9126-1 quality model came to be incorporated in the SQuaRE series (as ISO/IEC 25010), some ISO/IEC National Bodies commented on the discrepancy between the narrow definition of usability inherited from ISO/IEC 9126 and the broader definition in the CIF. But with the higher profile of usability in industry, there was now pressure to align the SQuaRE definition with the CIF, rather than vice versa. This was achieved by renaming the narrower ISO/IEC 9126 concept of usability as operability. This made it possible to define **Usability** as a characteristic of Quality in Use, with sub-characteristics of Effectiveness, Efficiency and Satisfaction. Quality in use in ISO/IEC CD 25010.3 has two further characteristics: **Safety** from 9126-1, and a new characteristic: **Flexibility**. These three characteristics are described in more detail below.

The complete quality in use model is:



2.1 Usability as a characteristic of quality in use

In the current draft of ISO/IEC CD 25010.3, usability has the same subcharacteristics as in ISO 9241-11. Two of the subcharacteristics retain the same definitions:

effectiveness (defined in terms of "accuracy and completeness") and **efficiency** (defined in terms of "resources expended").

But **satisfaction**, which is defined in ISO 9241-11 in terms of "comfort and acceptability of use", has been given a broader interpretation in ISO/IEC CD 25010.3. As Hassenzahl points out, current approaches to satisfaction typically assess primarily the users' perception of effectiveness and efficiency, so that if users perceive the product as effective and efficient, they are assumed to be satisfied [9]. But there is evidence that fun or enjoyment is an aspect of user experience that also contributes significantly to overall satisfaction with a product [8].

So in order to encompass the overall user experience, satisfaction needs to be concerned with both pragmatic and hedonic user goals.

The pragmatic user goals are:

- Acceptable perceived experience of use (pragmatic aspects including efficiency).
- Acceptable perceived results of use (including effectiveness).
- Acceptable perceived consequences of use (including safety).

In ISO/IEC CD 25010.3 it is suggested that these can be summarised as:

Likability (cognitive satisfaction): the extent to which the user is satisfied with the ease of use and the achievement of pragmatic goals, including acceptable perceived results of use.

Trust (satisfaction with security): the extent to which the user is satisfied that the product will behave as intended and with acceptable perceived consequences of use.

Hassenzahl identifies three hedonic goals [10]:

- Stimulation (i.e. personal growth, an increase of knowledge and skills).
- Identification (i.e. self-expression, interaction with relevant others).
- Evocation (i.e. self-maintenance, memories).

To these I would add:

- Pleasurable emotional reactions to the product (Norman's visceral category [23]).

In ISO/IEC CD 25010.3 it is suggested that these can be summarised as:

Pleasure (emotional satisfaction): the extent to which the user is satisfied with their perceived achievement of hedonic goals of stimulation, identification and evocation and associated emotional responses.

In addition ISO/IEC CD 25010.3 includes:

Comfort (physical satisfaction): the extent to which the user is satisfied with physical comfort.

Measuring satisfaction

Many software developers regard satisfaction as a personal response that cannot be quantified, and in much usability testing only qualitative feedback on satisfaction is obtained. Ad hoc questionnaires are sometime used, but psychometrically designed questionnaires will give more reliable results [12].

Simple questionnaires (such as SUS [5]) just measure the user's assessment of the ease of use. Longer questionnaires can measure more specific aspects, such as affect, efficiency, helpfulness, control and learnability in SUMI [20]. Trust can be measured using the System Trust Scale [19], and pleasure with questionnaires such as AttrakDiff [11]. There are also a variety of questionnaires for comfort, e.g. [24].

2.2 Flexibility as a characteristic of quality in use

Usability is defined in terms of user performance and satisfaction in a particular context of use. Considering context is important, as a product that is usable in one context of use may not be usable in another context with different users, tasks or environments. However, one consequence is that it may be possible to demonstrate very high usability for a product with carefully selected, but unrepresentative, users, tasks and environments of use. This was one of the motivations for the Common Industry Format: to enable a potential purchaser to judge whether the users, tasks and environments for which usability had been demonstrated, matched their own needs [25]. The Common Industry Specification for Usability – Requirements (CISU-R) went a step further, including a specification of the context of use in which usability is to be achieved as part of the usability requirements [4].

The need to consider the context has been made explicit in the new definition of quality in use in ISO/IEC CD 25010.3. A third characteristic of Flexibility has been included, with subcharacteristics of Context Conformity and Context Extendibility.

Context Conformity is defined as the degree to which usability and safety meet requirements in all the intended contexts of use. This provides the basis for measuring the extent to which usability has been achieved in the intended contexts of use.

Context Extendibility is defined as the degree of usability and safety in contexts beyond those initially intended. Context extendibility can be achieved by adapting a product for additional user groups, tasks and cultures. Context extendibility enables products to take account of circumstances, opportunities and individual preferences that may not have been anticipated in advance. If a product is not designed for context extendibility, it may not be safe to use the product in unintended contexts.

Accessibility is defined as the degree of usability for users with specified disabilities. This definition was used in preference to the ISO 9241-171 definition: "the usability of a product, service, environment or facility by people with the widest range of capabilities", as the ISO 9241-171 definition is stated as an objective that is difficult to quantify.

In an earlier draft, **learnability** was also a subcharacteristic of flexibility. Given the importance of setting objectives for learnability as usability in a learning context (effectiveness, efficiency and satisfaction with achieving learning goals), learnability could be reintroduced, for example as "the extent to which a product can be used by specified users to achieve specified learning goals with effectiveness, efficiency and satisfaction in a specified context of use". This would make learnability analogous to accessibility, with associated product characteristics for "technical learnability".

Measuring Flexibility

If usability requirements have been specified in a format similar to the CISU-R, testable usability objectives for user performance and satisfaction will have been given for specified ranges of context of use (user groups, tasks and environments). These requirements could in principle be evaluated in usability tests. In practice scarce testing resources will have to be prioritised, but identifying the range of intended contexts of use can reveal additional users, tasks and environments that may be a priority for testing.

It is important to specify requirements for flexibility, even if it is not practical to test them. The wider range of user groups, tasks and environments can be translated into additional design requirements for product features needed to support usage for all the identified contexts of use. Context Conformity in contexts that cannot be tested can be assessed by expert judgment.

Context Extendibility is more difficult to specify and measure in advance, as it is concerned with usage in unanticipated contexts of use. Design requirements that facilitate Context Extendibility include designing a product so that it can either be configured for specific needs (e.g. language, culture, task steps), or can be adapted by the user to suit individual capabilities and needs. Products are frequently used for unanticipated purposes (for example, use of Excel for prototyping [2]), so the ability to adapt the product to new needs significantly extends the usability. This can only be properly evaluated by patterns of actual usage, but the potential can be assessed by the extent to which the design is open to configuration and adaptation. As with Context Conformity, identifying requirements for Context Extendibility can have a large impact on design.

Accessibility can be specified by establishing objectives for usability for users with particular types of disabilities, thus giving a design objective for the success rate and productivity to be expected of users with disabilities. This can subsequently be validated by user testing.

2.3 Safety as a characteristic of quality in use

Safety is defined in ISO/IEC CD 25010.3 as the degree of expected impact of harm to people, business, data, software, property or the environment in the intended contexts of use.

While effectiveness and efficiency measure the positive benefits of productivity and goal achievement, the term safety is here interpreted in a broad way to measure the potential negative outcomes that could result from incomplete or incorrect output. For a consumer product or game, negative business consequences may not only be associated with poor performance, but also, for example, with a lack of pleasurable emotional reactions or of achievement of other hedonic goals.

Safety has four subcharacteristics:

- **Commercial damage**: the degree of expected impact of harm to commercial property, operations or reputation in the intended contexts of use. This could include the administrative costs of correcting erroneous output, inability to provide an acceptable service, or loss of current or future sales. Examples include lack of sales due to poor web site design, and the chaos following introduction of a new system to issue passports in the UK [25,21].
- **Operator health and safety**: the degree of expected impact of harm to the operator in the intended contexts of use. The legislation for workstation design in the EU [7] is intended to minimize this risk.
- **Public health and safety**: the degree of expected impact of harm to the public in the intended contexts of use.
- **Environmental harm**: the degree of expected impact of harm to property or the environment in the intended contexts of use.

Specification and measurement of usability should always be considered in conjunction with associated safety risks.

Measuring Safety

It is not easy to measure product safety, but it is usually possible to list the potential adverse consequences of product failure or human error, based on previous experience with similar products or systems. Multiplying the estimated frequency by the estimated impact of the potential failures provides a basis for prioritization of the need to find design solutions to minimize the probability of failures occurring. The solutions can include improving the user interface to reduce the probability of human error.

2.4 Quality in use for different stakeholders

One reason for the popularity of the ISO 9241-11 definition with usability professionals, is that when interpreted from the perspective of the organisation's goals it provides a business rationale for the importance of usability that is more compelling than mere ease of use.

But usability can also be seen from the inside out as meeting the user's goals, rather than the organisation's goals, which takes usability back closer to its original meaning. From this perspective the key element in the ISO definition is satisfaction, and this is one reason for the expanded interpretation of satisfaction in ISO/IEC 25010.3. For an officer worker and their manager cognitive satisfaction may be most important, for a games player or manufacturer, pleasure may be more important.

Measuring quality in use from different stakeholder perspectives

Usability and flexibility are measured by effectiveness (task goal completion), efficiency (resources used) and satisfaction. The relative importance of these measures depends on the purpose for which the product is being used (for example in some personal situations, resources may not be important).

Table 1 illustrates how the measures of effectiveness, resources, satisfaction, flexibility and safety can be selected to measure quality in use from the perspective of different stakeholders.

From an organisational perspective, quality in use and usability are about achievement of task goals. But for the end user there are not only pragmatic task-related "do" goals, but also hedonic "be" goals [6]. For the end user, effectiveness and efficiency are the pragmatic goals, and stimulation, identification, evocation and pleasure are the hedonic goals.

From an end user perspective, quality in use can be used to measure the extent to which users achieve their goals within an acceptable time, with hedonic and pragmatic satisfaction and without adverse consequences to the user's health and safety. They may also want to modify the interaction and appearance to suit their individual needs and preferences.

From the perspective of the organization using the product, the measures could be the extent to which the task goals are achieved with acceptable cost, with management satisfaction and with acceptable financial consequences to the business

of potential errors. They may also want to be able to customize the system to meet changing needs.

Table 1. Stakeholder perspectives of quality in use

STAKEHOLDER:	**End User** *Usability*	**Usage Organisation** *Cost-effectiveness*	**Technical support** *Maintenance*
GOAL: **MEASURES**	Pragmatic and hedonic goals	Task goals	Support goals
Effectiveness	User effectiveness	Task effectiveness	Support effectiveness
Resources	Productivity (time)	Cost efficiency (money)	Support cost
Satisfaction	Hedonic and pragmatic satisfaction	Management satisfaction	Support satisfaction
Flexibility	Individualisation	Customisation	Adaptability
Safety	Risk to user (health and safety)	Commercial risk	System failure or corruption

Similarly from a support perspective, the measures could be the extent to which the support goals, including adaptability, are achieved with acceptable cost, with support personnel satisfaction and with acceptable financial consequences to the business of any resulting human or software errors.

3 Measuring product attributes or quality in use

Table 2 shows how measures of usability and safety are dependent on product attributes that support different aspects of user experience. In Table 2 the columns are the quality characteristics that contribute to the overall user experience, with the associated product attributes needed to achieve these qualities.

The quality characteristics are frequently assessed by direct measurement of product attributes (the ISO/IEC 9126/25010 internal and external measures), for example by using heuristic evaluation, web design guidelines or the Web Content Accessibility Guidelines (WCAG) developed by the WAI [26]. However, these are only a means to an end: acceptable user performance and satisfaction.

The users' goals may be pragmatic (to be effective and efficient), and/or hedonic (stimulation, identification and/or evocation).

The actual experience of usage is difficult to measure directly. The measurable consequences are the user's performance, satisfaction with achieving pragmatic and hedonic goals, comfort and pleasure.

Table 2. Factors contributing to system usability and UX

Quality characteristic	Attrac-tiveness	Functional suitability	Ease of use	Learnability	Technical accessibility	Safety
Product attributes	Aesthetic attributes	Appropriate functions	Good UI design	Learnability attributes	Accessibility attributes	Safe and secure design
Pragmatic do goals			To be effective and efficient			
Hedonic be goals			Stimulation, identification and evocation			
Actual experience	Visceral		Experience of interaction			
Performance measures	Effectiveness and Efficiency: effective task completion and efficient use of time			Learnability: effective and efficient to learn	Accessibility: effective and efficient with disabilities	Safety: occurrence of unintended consequences
Satisfaction measures	Pleasure		Likability and Comfort			Trust

User performance and satisfaction is determined by qualities including attractiveness, functional suitability and ease of use. Other quality characteristics will also be relevant in determining whether the product is learnable, accessible, and safe in use.

Pleasure will be obtained from both achieving goals, and as a direct visceral reaction to attractive appearance [23].

Note that this framework does not include the popular measure of the number of errors made by the user, as the consequences of errors are already incorporated in effectiveness, efficiency and satisfaction. Uncorrected errors will impact on successful task completion, and corrected errors will contribute to task time and reduced satisfaction. However, counts of errors can be a useful indicator of usability problems, particularly when it is not practical to obtain measures of successful task completion.

4 Conclusions

The new definition of quality in use provides a framework for a more comprehensive approach to specifying usability requirements and measuring usability, taking account of the following issues.

1. From which stakeholder perspective(s) (e.g. users, staff and/or managers) does usability need to be specified and measured?
2. What is the scope of the context of use in which it is important to establish usability requirements?
3. What aspects of effectiveness, efficiency and satisfaction are most important, and how can they be measured?

4. What are the acceptable levels of risks of potential adverse consequences to the identified stakeholders resulting from poor usability or inappropriate output?
5. What product attributes are needed to achieve the identified objectives?
6. How can these product attributes be monitored during development?
7. Which are the most important contexts of use in which to validate usability?

References

1. ANSI: Common industry format for usability test reports (ANSI-NCITS 354-2001). ANSI (2001)
2. Berger, N., Arent, M., Arnowitz, J., Sampson, F.: Effective Prototyping with Excel: A practical handbook for developers and designers. Morgan Kaufmann (2009)
3. Bevan, N.: Quality in use: meeting user needs for quality, Journal of Systems and Software, 49(1), pp 89–96 (1999)
4. Bevan N, Claridge N, Maguire M, Athousaki M: Specifying and evaluating usability requirements using the Common Industry Format: Four case studies. Proceedings of IFIP 17th World Computer Congress, Montreal, Canada, 25-30 August 2002, 133-148. Kluwer Academic Publishers (2002)
5. Brooke, J.: SUS: A "quick and dirty" usability scale. In: Jordan, P., Thomas, B., Weerdmeester, B. (eds.) Usability Evaluation in Industry, pp. 189–194. Taylor and Francis (1996).
6. Carver, C. S., & Scheier, M. F.: On the self-regulation of behavior. Cambridge University Press (1998)
7. CEC: Minimum safety and health requirements for work with display screen equipment Directive (90/270/EEC). Official Journal of the European Communities No L 156, 21/6/90 (1990)
8. Cockton, G.: Putting Value into E-valu-ation. In: Law, E. L., Hvannberg, E. T., Cockton, G. (eds.) Maturing Usability. Quality in Software, Interaction and Value. Springer (2008)
9. Hassenzahl, M.: The effect of perceived hedonic quality on product appealingness. International Journal of Human-Computer Interaction, 13, 479–497 (2002)
10. Hassenzahl, M.: The thing and I: Understanding the relationship between user and product. In: Blythe, M., Overbeeke, C., Monk, A.F., Wright, P.C. (eds) Funology: From Usability to Enjoyment, pp. 31–42. Kluwer, Dordrecht (2003)
11. Hassenzahl, M.: AttrakDiff(tm), www.attrakdiff.de
12. Hornbæk, K.: Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer Studies, 64: 79–102 (2006)
13. ISO/IEC 9126: Software engineering – Product quality. ISO (1991)
14. ISO/IEC 9126-1: Software engineering – Product quality - Part 1: Quality model. ISO (2001)
15. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) Part 11: Guidance on Usability. ISO (1998)
16. ISO 9241-171: Ergonomics of human-system interaction – Part 171: Guidance on software accessibility. ISO (2008)
17. ISO/IEC CD 25010.3: Systems and software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Software product quality and system quality in use models. ISO (2009)
18. ISO/IEC 25062: Software Engineering - Software product Quality Requirements and Evaluation (SQuaRE)-Common Industry Format (CIF) for Usability Test Reports. ISO (2006)

19. Jian, J-Y., Bisantz, A. M., Drury, C. G: Foundations for an empirically determined scale of trust in automated systems. International Journal of cognitive Ergonomics, 4, 1, 53-71 (2000)
20. Kirakowski, J.: The Software Usability Measurement Inventory: Background and Usage. In: Jordan, P., Thomas, B., Weerdmeester, B. (eds) Usability Evaluation in Industry. Taylor and Francis (1996)
21. National Audit Office: United Kingdom Passport Agency: The passport delays of Summer 1999. Publication HC 812 1998-1999 (1999)
22. Nielsen, J.: Usability Engineering. Academic Press (1993)
23. Norman, D.: Emotional design: Why we love (or hate) everyday things. New York: Basic Books (2004)
24. Norman K, Alm H, Wigaeus Tornqvist E, Toomingas A.: Reliability of a questionnaire and an ergonomic checklist for assessing working conditions and health at call centres. Int J Occup Saf Ergon,12:53–68 (2006)
25. Theofanos, M., Stanton, B., Bevan, N.: A practical guide to the CIF: Usability Measurements. Interactions 18 (6) 34-37 (2006)
26. Web Accessibility Initiative (WAI): Web Content Accessibility Guidelines 2.0, http://www.w3.org/TR/WCAG20/