Applying the ISO 9126 model to the evaluation of an e-learning system

Bee Bee Chua and Laurel Evelyn Dyson
Faculty of Information Technology
University of Technology, Sydney, AUSTRALIA

Despite the widespread use of e-learning systems and the considerable investment in purchasing or developing them in house, there is no consensus on a standard framework for evaluating system quality. This paper proposes the ISO 9126 Quality Model as a useful tool for evaluating such systems, particularly for teachers and educational administrators. The authors demonstrate the validity of the model in a case study in which they apply it to a commonly available e-learning system and show how it can be used to detect design flaws. It is proposed that the metric would be applicable to other e-learning systems and could be used as the basis for a comparison to inform purchase decisions.

Keywords: e-learning, ISO, ISO 9126, Blackboard, online learning

Introduction

Most universities and colleges use e-learning systems to support face to face learning in the classroom or to implement distance learning programmes. The growth of e-learning systems has increased greatly in recent years thanks to the demand by students for more flexible learning options and economic pressures on educational institutions, who see technology as a cost saving measure. Yet, there has been considerable criticism of the quality of the systems currently being used. Problems include low performance, poor usability, and poor customisability, which make it difficult to serve the specific needs of different learners. Furthermore, online education has often been criticised as not supporting learner centred education but replicating traditional face to face instruction (Vrasidas 2004).

Despite the widespread use of e-learning systems and the considerable investment in purchasing or developing them in house, there is no consensus on devising a standard framework for evaluating system quality in this area. The lack of an agreed e-learning system quality model is in stark contrast to the extensive work on software quality assurance in general (Crosby 1979; Garvin 1984; Juran 1988; Norman & Pfleeger 2002).

This paper proposes the ISO 9126 Quality Model (ISO 1991) as a useful tool for evaluating such systems. The ISO 9126 model was developed by the International Organization for Standardisation (ISO) and is one of a large group of internationally recognised standards applicable across a wide range of applications. To date, ISO 9126 has not been applied extensively to the e-learning environment. Nevertheless, the authors believe that it has potential to provide a useful evaluation tool: this belief derives from the many years of industry experience that one of the researchers has had in software quality assurance. Perspectives from this domain could provide insights relevant to e-learning educators. In this paper we propose that the ISO 9126 model could be used as the basis for a comparison of e-learning systems to inform decisions regarding review of existing systems and the purchase of new ones.

First of all, the paper examines the e-learning system literature and evaluates some of the software quality tools and frameworks that have been proposed. Secondly, we introduce the ISO 9126 Quality Model as a basis for evaluating e-learning tools and explain the characteristics and sub-characteristics of the model. The main objective of our paper was to demonstrate how the model can be used to evaluate an e-learning system. With this in mind, we chose a commonly used system, Blackboard, as a basis for our research and adopted a case study approach. We applied the model to the system in the context of an Information Technology subject in an undergraduate programme. In this paper, we summarise the results of the evaluation of the system: generally, our results show the model is a good framework for assessing e-learning systems, although we do identify several possible refinements to the model. Finally, we analyse the implications of using the ISO 9126 Quality Model to evaluate and improve e-learning systems.

E-learning system quality

Research into evaluating e-learning systems comes from two directions: the educationalists and the software developers. Many educators have shown significant interest in the pedagogical evaluation of e-learning systems, that is, in course design issues and how to promote good learning (Laurillard 1993; Reeves 1992). Although these fundamental issues of course design are vital, these studies do not assist educators in evaluating the quality of the *system* as such, and therefore do not incorporate frameworks to support decision making regarding review of existing systems and the purchase of new ones.

There is also a vast body of literature relating to various technical frameworks for software developers who wish to improve the quality of the e-learning systems they are developing. A systematic approach is the IEEE Learning Technology Standard Committee (LTSC) reference model, IEEE P1484.1 LTSA. This model has five layers, which focus on reusability and portability, and compares different e-learning systems by numerical rating scales for various factors, e.g., assessment, administration, curriculum development, etc. (O'Droma, M. S., Ganchev, I. & McDonnell, F. 2003). The Sharable Content Object Reference Model (SCORM) is another widely known framework. It supports content compatibility, that is the portability of content from one e-learning system to another and the re-usability of learning objects by extensive cataloguing using metadata (Bohl, Schelhase, Sengler & Winand 2002). The Instructional Management Systems (IMS) project is another approach to defining technical specifications in order to promote interoperability between e-learning systems (IMS Global Learning Consortium). These standards focus on technical aspects of e-learning systems and neglect the Human Computer Interaction (HCI) component, that is, how the user will interact with the system. More importantly, they are too complicated for the average educator or educational administrator to understand and apply when choosing an e-learning system. They are specially designed for technical trained system developers.

The few studies that have been undertaken for educators and people working in educational institutions who need to evaluate e-learning systems are often inadequate. This is due to the lack of systematic tools or approaches. For example, Roberts (2002) gains a good overview of Blackboard using surveys, focus groups and interviews, but the results are too general and do not provide detailed analysis of features such as usability. In another study, the Learning and Teaching Technology Group (LTTG) undertook a comparison of Blackboard and WebCT. Their main approach was an evaluation based on the number of times students accessed different tools in the system, e.g. discussion board, group areas and others. Nevertheless, these quantitative counts are not meaningful without details of the subject design, for example how the group area activities were incorporated into the learning environment. The rest of their paper offers a miscellaneous group of features for evaluation, e.g. data integration, pricing, hardware or software platforms and ease of access. However, there is no system or justification for their choice of features and many common usability criteria are omitted. A third approach we examined was "20 Questions", which Driscoll and Dennehy (2002) propose putting to suppliers of the system. They resolve the adoption of an e-learning system into two factors, organisational and technical, although only a few of their questions deal with organisational issues and the main emphasis is on the technical issues, e.g. back end integration and the partitioning of the system. Student interaction with the system receives very little attention in their approach. Likewise, Parisotto (2003) focuses broadly on high level issues in evaluating e-learning systems. He considers three organisational perspectives (academic, administrative and IT support) but fails to discuss the operational levels, that is, the system in use.

The ISO 9126 model

The International Organization for Standardisation (ISO) was founded in 1946 in order to facilitate international trade, international coordination and unification of industrial standards by providing a single set of standards that would be recognised and respected (Praxiom Research Group). ISO 9126 was originally developed in 1991 to provide a framework for evaluating software quality and then refined over a further ten year period (Abran et al. 2003). Many studies criticise ISO 9126 for not prescribing specific quality requirements, but instead defining a general framework for the evaluation of software quality (Valenti 2002). We believe that this is in fact one of its strengths as it is more adaptable and can be used across many systems, including e-learning systems. The original model defined six product characteristics (see Figure 1). These six characteristics are further subdivided into a number of sub-characteristics (see Table 1).

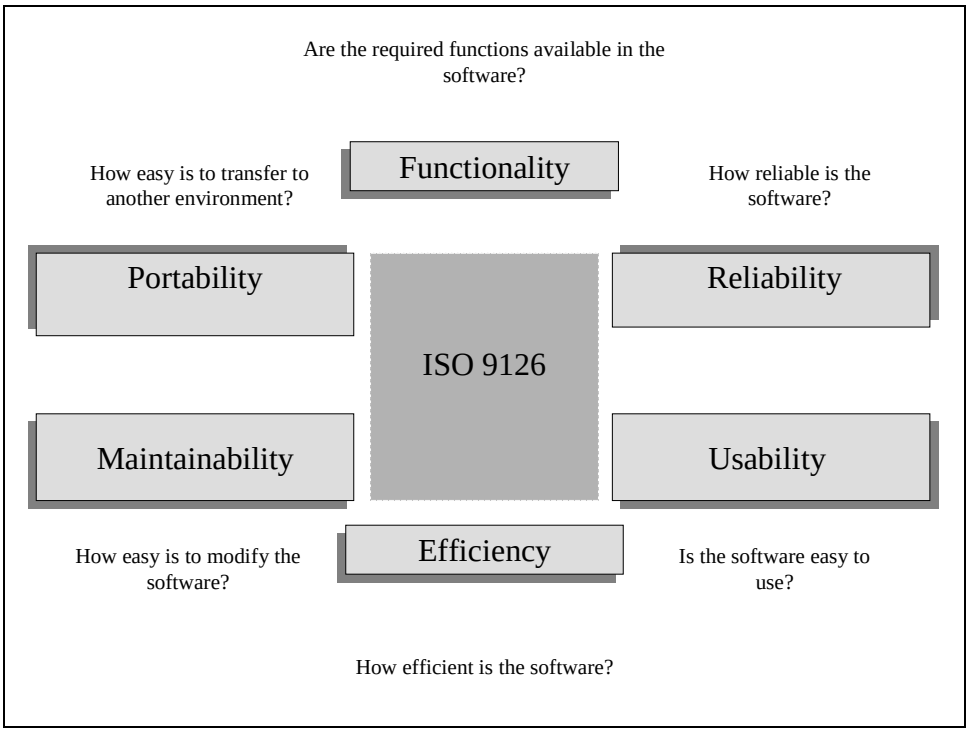

Figure 1: (Source: ISO 1991)

Table 1: ISO 9126 Characteristic and sub-characteristics (Source: ISO 1991; Abran 2003)

Characteristic	Sub-characteristic	Explanation
Functionality	Suitability	Can software perform the tasks required?
	Accurateness	Is the result as expected?
	Interoperability	Can the system interact with another system?
	Security	Does the software prevent unauthorised access?
Reliability	Maturity	Have most of the faults in the software been eliminated over time?
	Fault tolerance	Is the software capable of handling errors?
	Recoverability	Can the software resume working and restore lost data after failure?
Usability	Understandability	Does the user comprehend how to use the system easily?
	Learnability	Can the user learn to use the system easily?
	Operability	Can the user use the system without much effort?
	Attractiveness	Does the interface look good?
Efficiency	Time Behaviour	How quickly does the system respond?
	Resource Utilisation	Does the system utilise resources efficiently?
Maintainability	Analysability	Can faults be easily diagnosed?
	Changeability	Can the software be easily modified?
	Stability	Can the software continue functioning if changes are made?
	Testability	Can the software be tested easily?
Portability	Adaptability	Can the software be moved to other environments?
	Installability	Can the software be installed easily?
	Conformance	Does the software comply with portability standards?
	Replaceability	Can the software easily replace other software?
All characteristics	Compliance	Does the software comply with laws or regulations?

These characteristics and sub-characteristics represent a detailed model for evaluating any software system. Indeed, Abran, Khelifi, Suryn & Seffah (2003) claimed that, "Even though it is not exhaustive, this series constitutes the most extensive software quality model developed to date." It is also an easy model for the non-specialist to employ, for example, simpler than the IEEE P1484.1 LTSA model, SCORM or IMS. Unlike these other frameworks, ISO 9126 covers a wide spectrum of system features,

including both technical requirements and human interaction with the system. For example, ISO 9126 includes HCI features such as attractiveness of the interface, which is overlooked by the other standards.

Methodology

The researchers used the quality characteristics and sub-characteristics to evaluate an e-learning system, Blackboard version 6.1. From the educator's point of view, the first three characteristics (Functionality, Reliability and Usability) and the first sub-characteristic of Efficiency (Time Behaviour) are easily assessable, whereas the remaining characteristics are difficult to measure except by trained IT professionals (Valenti et al. 2002). For this reason, our focus will be on these earlier characteristics.

The evaluation centred on the use of the Blackboard system by students and teaching staff during one subject for one semester. The subject was being taught in a faculty of Information Technology and students had some experience in using the system in the previous semester. The students used the system both in a classroom environment and in their own time.

In our investigation several evaluation methods were employed. Firstly, we focused on the system in use by observing students while we were teaching them during semester. Secondly, our own experiences as teachers using the system were recorded. Thirdly, students and teacher contributions to discussion boards and group spaces within the e-learning system were examined as evidence of activity. Fourthly, we ran a test of the different tools within the system based on the characteristics and sub-characteristics of the ISO 9126 model including timing, fault detection and general usability and functionality. In general, the evaluation was qualitative although, for evaluating the subcharacteristic Time Behaviour, a time test was conducted of the system's performance on two different computers, one older and one a newer, faster machine, both operating on a fast Ethernet network, which has a bandwidth of 100 Kbps. This time test supplemented observation in class when up to 120 students were using the system simultaneously.

Results

The results were summarised into a matrix (adapted from Abel and Rout (1993)) relating the characteristics and sub-characteristics to the main tools offered by the e-learning system (see Figure 2). An asterisk in the matrix indicates that the tool satisfies the requirements of the sub-characteristic. Where deficiencies were identified in the evaluation, these have been indicated by a number and an explanation is given in the legend below of how the system failed to meet all the criteria in these cases.

Discussion

From our evaluation, we discovered many flaws with the system. Some of these are critical to user satisfaction and some are minor. This depends on who the user is (subject co-ordinator, teacher or student). The ISO 9126 model provides an indication to educators and educational administrators of the quality of a system they are considering buying into and provides a basis of comparison of different systems.

Though our results demonstrate the ISO 9126 model is useful in evaluating e-learning systems, the researchers also have some recommendations on how it could be enhanced. Firstly, we believe that it could be improved by having a global characteristic to summarise the overall user satisfaction. To determine the user satisfaction level, it is not possible to simply add up the number of problem sub-characteristics. Different users will have different priorities that will influence on which characteristics they will place more emphasis. Therefore, we need to consider incorporating a final characteristic for the user to state whether the particular tool being evaluated is acceptable overall or not.

Secondly, the sub-characteristic Appearance is too general and covers too many different factors and is therefore not very helpful. It is recommended that the sub-characteristics included under Usability be extended to include more specific appearance factors based on accepted HCI usability principles. For example, Usability should include the sub-characteristics consistency, simplicity, legibility (e.g font size) and use of colour (Preece, Rogers & Sharp, 2002).

TOOLS	Quality characteristics											
	Functionality				Reliability			Usability				Efficiency
	Suitability	Accurateness	Interoperability	Security	Maturity	Fault Tolerance	Recoverability	Understandability	Learnability	Operability	Attractiveness	Time Behaviour
Students' and Teacher's tools:												
Course announcements	*	*	*	*	*	*	*	*	*	*	*	*
Course information	*	*	*	*	*	1	2	2	2	*	3	*
Information about teachers	*	*	*	*	*	1,4	2	2	2	*	3	*
Teaching materials	*	*	*	*	*	1,5	2	2	2	*	3	*
Assignments	*	*	*	*	*	*	*	*	*	*	3	*
Class discussion board	*	*	*	*	*	*	*	6	6	6	*	*
Group discussion board	*	*	*	*	*	*	*	6	6	6	*	7
Class chat room	*	*	*	*	*	*	*	8	8	*	3	9
Group chat room	*	*	*	*	*	*	*	8	8	*	3	7,9
Class drawing tool	10	*	*	11	*	*	*	12	12	*	*	9,13
Group drawing tool	10	*	*	*	*	*	*	12	12	*	*	7,9
Roster	14	*	14	*	*	*	*	14	14	*	3	*
Email	*	*	*	*	*	*	*	*	*	*	3	*
Group file exchange	*	*	*	*	*	*	*	*	*	*	3	7
Calendar	*	*	*	*	*	*	*	*	*	*	*	*
Teacher's tools:												
Manage groups	15	*	*	*	*	16	16	15	15	17	3,18	17
Grade book	*	*	*	*	*	*	*	*	*	*	*	*
Tests	*	*	*	*	*	*	*	*	*	*	19	*
Course statistics	*	*	*	*	*	*	*	*	*	*	*	*

Figure 2: Evaluation of e-learning using ISO 9126

Matrix Legend:

1. Accepts null content when content is logically required.
2. Lack of labeling regarding the required field makes it more difficult to use.
3. Font size is too small. Huge inconsistencies in font from one page to another.
4. When uploading a non-standard size picture of the staff member, an incorrect message was displayed.
5. The system does not check for validity of dates when teaching materials will become available.
6. Poor navigation. A menu of navigation buttons is needed instead of the one button provided, and these need to be clearly named according to their function.
7. Loading group page was very slow when lots of users were online.
8. Problem with interpreting non-standard terminology, for example, "virtual classroom", "room available in the future".
9. Chat room is very slow in initialising due to a need to install Java Applet plug in.
10. Can't save drawings within the system or export drawings.
11. Because of anonymity of drawing and graphics upload there was a problem with a pornographic image being posted.
12. The function of all buttons was not easy to understand. Tool tips are needed.
13. There is a synchronicity problem, with a time lag between when students can see what fellow student have drawn.
14. Poor functionality and hard to understand how to use it: unable to display a roster.
15. Cannot search on user's first name and cannot list all group members. 'List' button is therefore hard to understand.
16. When creating new groups, system is unable to cope with a too long group description.
17. Adding students to a group involves seven mouse clicks from one side of the screen to another *for every single student.* Most of the buttons involved cannot be activated from the keyboard. This impacts time behaviour.
18. Order of groups is not alphabetical.
19. Inconsistent layout.

Thirdly, there needs to be a way of evaluating help given to users in the system. Currently, the ISO model does not include any way of explicitly evaluating this Usability principle. For example, we had no way of recording a major deficiency with the system we evaluated, the fact that there was no help for students apart from very limited tool tips on some screens only and a login help. The researchers recommend that the model include a Help sub-characteristic under Usability.

Fourthly, a strong correlation was found between the sub-characteristics Understandability and Learnability. Therefore, we recommend a combination of these two sub-characteristics into one. Lastly, Maturity was found to apply to the system as a whole but is not useful in evaluating individual tools: for example, we know the system is mature (version 6.1) but we have no way of knowing the maturity of each tool.

Conclusion

In this paper, we have discussed how the ISO 9126 Quality Model can be applied to evaluating e-learning systems. It provides a detailed analytical tool and is useful in moving beyond superficial evaluation to achieve a more thorough view of the system's strengths and weaknesses than can be provided by less systematic approaches. For teachers and administrators in the educational field, who need to make decisions regarding which system to buy, it provides a possible metric for comparison of the various products available on the market. As such it can provide a basis for informed and rational decision making and avoid costly mistakes.

However, in our investigation we uncovered some inherent weaknesses in the model, particularly with regards to the Usability characteristic. To make the model simpler to use for educators, who may not be usability experts, we propose that this characteristic should be extended to include more specific factors such as consistency, simplicity, legibility and colour use. It is also suggested that a Help sub-characteristic be included as part of Usability, mainly to ensure that this important factor should not be neglected. In addition, we propose the inclusion of user satisfaction as a global characteristic to summarise the general impact of the system on the user in their specific educational context and given their specific requirements. With these improvements, ISO 9126 could be a useful model for evaluating the quality of e-learning systems.

A question that must be addressed is whether ISO 9126 could be used by software development companies who are interested in developing new e-learning systems. Could this model provide a testing benchmark for software quality evaluation of the product prototypes and eventually lead to better final products with enhanced user satisfaction? It must be said that the success of our application of ISO 9126 depended not only on the strength of the model but also on the researchers' teaching expertise as experienced and qualified teachers utilising the e-learning system in an educational environment: we knew the various pedagogical functions that the system had to support. However, for software developers without educational expertise, the ISO model alone would be insufficient because it is a general software quality model and does not specify the particular teaching and learning activities needed for good learning. For this group, a checklist of tools and attributes which promote good educational outcomes and efficient course management would be needed, such as those proposed by Britain and Liber (2003).

Further to our current study, we intend to obtain feedback from teachers and students, for example using the following research methods: student questionnaires, teacher interviews and focus groups. This will help cross validate the usefulness of the ISO 9126 Model for evaluating e-learning software quality. This further research should motivate educators to perceive the benefits of having a standard to underpin needed improvements in the e-learning context, particularly as these systems move into cutting edge, multimedia technologies in the near future.

References

Abel, D. E., & Rout, T. P. (1993). The Mapping of Software Quality Engineering to ISO 9126. *Australian Computer Journal*, 25(August), 1-14.

Abran, A., Khelifi, A., Suryn, W., & Seffah, A. (2003). Usability Meanings and Interpretations in ISO Standards. *Software Quality Journal*, 11(4), 325-338.

Bohl, O., Schelhase, J, Sengler, R. & Winand, U. (2002). The Sharable Content Object Reference Model (SCORM) – A Critical Review. *Proceedings of the International Conference on Computers in Education (ICCE'02).*

Britain, S. & Liber, O. (2003). A Framework for Pedagogical Evaluation of Virtual Learning Environments. [verified 10 Oct 2004] http://www.leeds.ac.uk/educol/documents/00001237.htm

Crosby, P.B. (1979). *Quality Is Free: The Art of Making Quality Certain.* McGraw-Hill, New York.

Driscoll, M. & Dennehy, M. (2002). Playing 20 Questions. [verified 10 Oct 2004] http://www.learningcircuits.org/2002/dec2002/driscoll.htm

Garvin, D.A. (1984). What does 'Product Quality' really mean? *The Sloan Management Review*, 26(1), 25-43.

IMS (Instructional Management Systems) (2002). IMS Frequently Asked Questions. IMS Global Learning Consortium, Inc. http://www.imsglobal.org/faqs/imsnewpage.cfm?number=2

International Organization for Standardisation. (ISO) (1991). ISO/IEC: 9126 Information technology-Software Product Evaluation-Quality characteristics and guidelines for their use –1991. [verified 10 Oct 2004] http://www.cse.dcu.ie/essiscope/sm2/9126ref.html

Juran, J. M. (1988). *Juran on Planning for Quality*. Free Press, New York.

Laurillard. D, (1993). *Rethinking University Teaching*. Routledge, London.

LTTG (Learning and Teaching Technology Group), London School of Economics. *Blackboard 5.0 Evaluation*. [verified 10 Oct 2004] http://teaching.lse.ac.uk/tech/blackboardeval/

Martin, M. & Jennings, A. (2002). *eLearning Technology Evaluation Report*. [verified 10 Oct 2004] http://www.elearningalliance.org/uploads/attachments/eLearning%20Technology%20Evaluation%20Report.pdf

Norman, E. F. & Pfleeger, L. S. (2002). *A Rigorous and Practical Approach*. 2nd Edition. PWS Publishing Company, London, Boston.

O'Droma, M. S., Ganchev, I. & McDonnell, F. (2003). Architectural and functional design and evaluation of e-learning VUIS based on the proposed IEEE LTSA reference model. *The Internet and Higher Education*, 6(3), 263-276. [verified 10 Oct 2004] http://www.ul.ie/~odroma/Architectural%20and%20functional%20%20Internet%20and%20HE%202003.pdf

Parisotto, L. (2003). BCIT CMS Evaluation Plan. [verified 10 Oct 2004] http://online.bcit.ca/sidebars/03september/inside-out-2.pdf

Preece, Rogers & Sharp (2002). *Interaction Design Beyond Human-Computer Interaction.* John Wiley & Sons, Inc, USA.

Roberts, G. (2002). An evaluation of use of Blackboard VLE in first year, core business modules, 2001-2002. [verified 10 Oct 2004] http://www.rgu.ac.uk/files/Gillians%20Paper.doc

Reece, G. A. (2002). ADA–508 Compliance and Usability Testing Techniques for Accessible Web Pages. [verified 10 Oct 2004] http://www.stc.org/confproceed/2002/PDFs/STC49-00001.pdf

Reeves, T. C.(1992). Effective dimensions of interactive learning systems. Invited keynote paper presented at the Information Technology for Training and Education Conference (ITTE'92), Queensland, Australia.

Valenti, S., Cucchiarelli, A. & Panti, M. (2002). Computer Based Assessment Systems Evaluation via the ISO9126 Quality Model. *Journal of Information Technology Education*, 1(3), 157-175.

Vrasidas,C. (2004). Issues of Pedagogy and Design in e-learning Systems. *2004 ACM Symposium on Applied Computing*, pp. 911-915.

Bee Bee Chua can be contacted on bbchua@it.uts.edu.au
Laurel Evelyn Dyson can be contacted on laurel@it.uts.edu.au

Please cite as: Chua, B.B. & Dyson, L.E. (2004). Applying the ISO9126 model to the evaluation of an e-learning system. In R. Atkinson, C. McBeath, D. Jonas-Dwyer & R. Phillips (Eds), *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 184-190). Perth, 5-8 December. http://www.ascilite.org.au/conferences/perth04/procs/chua.html