

Identifying Appropriate E-learning Usability Evaluators

Camille Dickson-Deane, University of Missouri, cdickson-deane@mizzou.edu

Joi L. Moore, University of Missouri, moorejoi@missouri.edu

Krista Galyen, University of Missouri, kdgrz9@mizzou.edu

Weichao Chen, University of Missouri, wcxcf@mizzou.edu

Ngoc Vo, University of Missouri, ntv25c@mizzou.edu

Margarida Washburn, University of Missouri, mawzt2@mizzou.edu

Abstract

Evaluators are normally selected based on their level of expertise within a particular field. When evaluating e-learning products the skills required can be blurred thus reducing the quality of the evaluation report. This pilot illustrates as a usability evaluation project is conducted, the evaluators may have difficulty understanding the usability instrument thus skewing the results of the evaluation process. This pilot has implications for usability studies, research instruments and the quality of evaluators selected for usability evaluations.

Keywords: e-learning, usability, evaluator

INTRODUCTION

The field of distance education has a long history which reflects the technology of the era in which it is used. Today, distance education takes a more popular form where ease of access and cost are the basis for the decisions in implementation. Technology has improved and with it, increased access has made electronic learning, normally abbreviated as e-learning, internationally accepted. In addition to this, a shift in attitude by not only practitioners but learners, along with the realization that the use of e-learning will allow greater dissemination of knowledge to the masses has increased the growth of e-learning. With acceptance, comes variations of models and offerings all bundled under the title of e-learning. The Sloan Consortium (Allen and Seaman, 2007) has attempted to solve this dilemma by identifying four types of courses; 1) Traditional course which has no content delivered online; 2) Web facilitated courses which have 1 to 29% of its content delivered online; 3) Blended/Hybrid courses which have 30 to 79% of its content delivered online and 4) Online course where there are normally no face to face instruction, 80% or more of its content is delivered online. The varying types of courses have, as a by-product, revealed a great concern regarding the quality of the products being provided. Evaluation of these products provides a huge area for quality management. Thorough investigations need to be undertaken when looking at each type of course, its delivery methodology, its objectives and its audience.

E-learning products have a various number of characteristics and based on the history of distance education a further variety of definitions for terms associated with the field. Similarly, instruments for usability evaluation differ based on the unit of analysis, i.e. course, program, learning object, etc... (Moore, Dickson-Deane, Galyen, Vo, and Charoentham, 2008). Most usability protocols expect evaluators to review e-learning products from a variety of views i.e. as the learner, as the instructor or even as an expert in interface or instructional design. Kirmani & Rajasekaran (2007) found when classifying evaluators into two levels of expertise based on the individual ability of the evaluator, certain usability interaction (UI) parameters had a greater quality rating than other parameters. These ratings were then greatly reduced when averaged across 18 evaluators with UI evaluation experience. The evaluators in these cases, reviewed web sites and web applications and not e-learning products. This

paper illustrates that in order to review the usability of e-learning products, the characteristics of the evaluators need to be identified/specified in order to provide a quality review.

METHOD

Team

Using a team comprised of faculty and doctoral students in the field of learning technologies, investigations in e-learning usability focused the usability evaluation of e-learning products. In order to receive quality usability reports, elements influencing the validation of usability evaluation were identified. These elements, the instrument, the evaluator, the evaluation process and the evaluation context, further produced the following questions:

- How do these four elements interact with each other?
- How do the identified interactions influence the quality of the evaluation report?
- Which factors belonging to the elements contribute to these interactions?

This encouraged further investigation into the role of the evaluator and hence the characteristics associated with the evaluator

- What knowledge is needed for the success of a usability evaluation?
- How do we determine evaluation expertise?

These final two questions were the focus of this pilot study.

Instrument description

Dringus and Cohen (2005) designed an instrument based on a list of usability problems for WebCT, a learning management system (LMS). The instrument is framed as an heuristic evaluation that includes both the learner and the instructor perspective. The original instrument contained a 161 questions in 13 categories. The researchers included a yes/no component to facilitate the rating process of a course.

Course description

An e-course, defined as a course offered with a LMS that combined several communication technologies, which is supported by instructor-led activities was used (Moore et al., 2008). The course topic was Adobe Flash and was offered as a 16 week (semester long) course. The course used, Sakai, an open source software and implemented navigational media, webpages and word documents to provide information on the required topics. This mimics in some aspects, the course that was used by Dringus and Cohen (2005).

Participants

Dreyfus and Dreyfus (1986) identified five levels of skill acquisition. In their classification, an acquirer begins as a novice, then progresses on towards an advanced beginner, then to competent, proficient and finally the expert level. Acquirers can be classified based on their own perception and decision-making abilities as opposed to routinized actions (Eraut, 1994). Skill here is defined as an "integrative overarching approach to professional action, which incorporates both routines and decision to use them." Four evaluators completed a demographic survey to illustrate their different skills used to review the course. This survey produced four specific skills for usability evaluation 1) abilities in Learning Management System (LMS), 2) online instruction, 3) instructional design and 4) usability evaluation. The survey was used to illustrate the level of expertise of each evaluator as per the Dreyfus and Dreyfus Model of Skills Acquisition (see Figure 1).

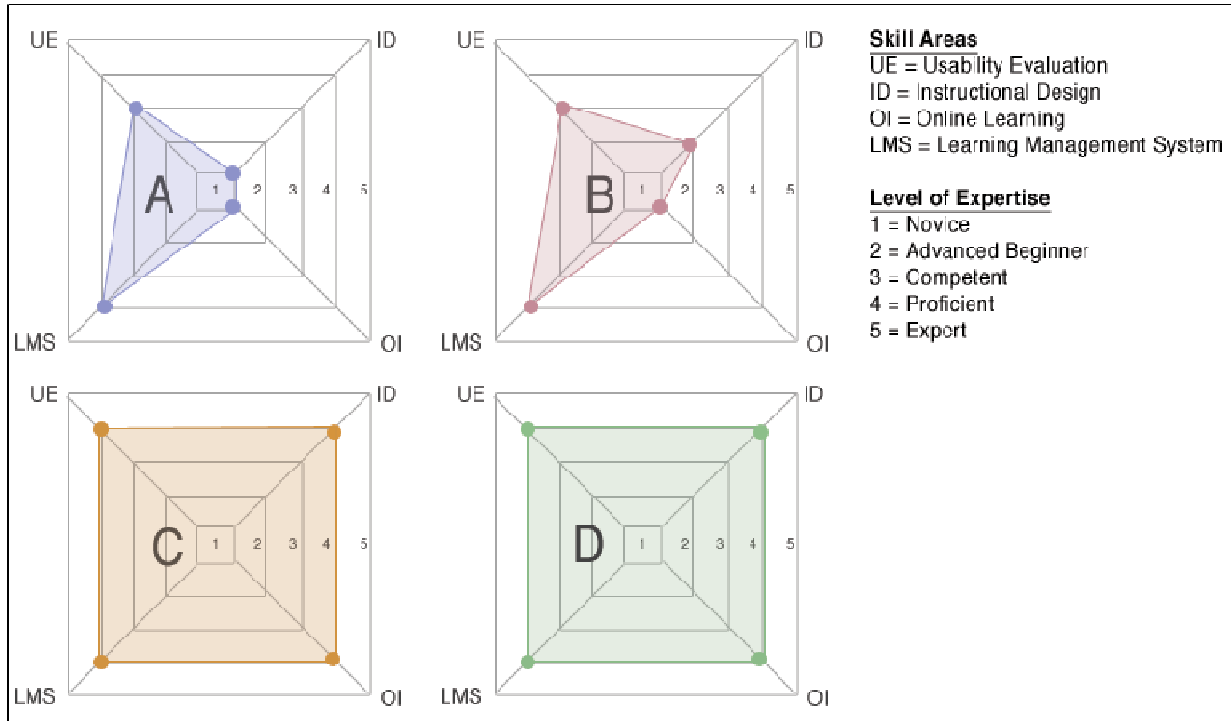


Figure 1 Evaluators mapped to Dreyfus and Dreyfus Skills Acquisition Model. (Total area is symbolic of overall expertise.)

Method

Using the Dringus and Cohen (2005) instrument as the anchor, we had five steps of evaluation:

1. An initial group-review of each category for inter-rater reliability was completed. This ensures consistency in application based on the meaning of the categories.
2. Each participant completed a demographic questionnaire
3. Each participant applied the instrument to the course on their own.
4. Each evaluator wrote a reflection of their evaluative actions
5. All the evaluators met as a group to reflect/share their interactions with the instrument and the course where the reflective session was recorded and transcribed.

The time to evaluate the course was not constrained. The results of the evaluation reports and the recorded reflections were analyzed and compared with the demographics of each evaluator.

FINDINGS

During Self and peer evaluations, several issues emerged:

- **Evaluation criteria-** Continued inconsistencies between the evaluator's understanding and application of different ratings for a category. In many cases, there were several situated responses, meaning it could not be a definitive yes/no response.
- **Clarity of the wording/phrasing of checklist items-** some of the checklist items were confusing to the evaluators. The words used in evaluation items were ambiguous in meaning e.g. appropriate, accessibility standards.
- **Generalization of LMS** - There were items that were more specific to particular types of LMSs e.g. Message can be easily expanded and collapsed. Sakai uses J Forums and hence does not have threaded discussions.
- **Evaluative perspective** - the checklist items mixed both student and instructor perspectives. Some evaluators took classes but never taught. The terminology suggests that an evaluator will be able to switch perspectives when using the checklist.
- **Knowledge for evaluation** - There was a challenge in interpreting categories, which is also based on the evaluator's experience. Instructional design questions and accessibility standards questions were present thus assuming that evaluators had both
- **Knowledge of evaluation methods** - Understanding usability evaluation did not arise as an issue but more so as a concern of the evaluators. The possibility that an evaluator may not have the knowledge of conducting usability evaluations can hinder the successful completion of the procedure.

CONCLUSION

The evaluators' reflective actions provided four key points of references for future evaluation considerations. A usability evaluation experience assumes that the evaluator has some foundational knowledge on usability standards. This large assumption hinges on the premise that the evaluation team may or may not consider foundational knowledge to be important enough for an evaluation team member. The second point elaborates on the first in that general as well as specific knowledge is needed to implement evaluation instruments. An instrument that uses key categories such as instructional design and interface design within its format are assuming that the users have the knowledge of these categories to review the product. The third and fourth points are more related to building stronger evaluators and hence stronger evaluative reports. Continuous self-reflections during the evaluative process to rationalize the decisions taken whilst using the instrument can build evaluation competencies. Finally, in the same note, group reflections can be used to synthesize and construct new views of evaluation through discussions. These reflections assist with building stronger skill sets for the evaluation process, and assist with the improvement of the product.

REFERENCES

Allen. I. E., and Seaman, J. (2007). Online Nation: Five years of Growth in Online Learning. *The Sloan*

Consortium, Olin Way, Needham, MA. Available at http://wiseplus.exp.sis.pitt.edu:8080/dspace/bitstream/123456789/47/1/online_nation.pdf [Accessed 19th February, 2009]

Dreyfus, H. L. and Dreyfus, S. E. (1986). *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. Oxford, Basil Blackwell.

Dringus, L. P., and Cohen, M. S. (2005). "An adaptable usability heuristic checklist for online courses". *Frontiers in Education*, 2005. FIE'05. Proceedings 35Th Annual Conference, T2H-6

Eraut, M. (1994). *Developing professional knowledge and competence*. London: Falmer

Kirmani, S. and Rajasekaran, S. (2007). Heuristic Evaluation Quality Score (HEQS): a measure of heuristic evaluation skills. *Journal of Usability Studies*, 2(2), 61-75.

Moore, J. L., Dickson-Deane, C., Galyen, K., Vo, N., and Charoentham, M. (2008). "e-Learning Usability Instruments: What is being Evaluated?" *Proceedings from E-Learn 2008*: Association for the Advancement of Computing in Education.