Usability Evaluation of a Nuclear Math and Theory Online Learning Environment

Krista Galyen, Joi L. Moore, Camille Dickson-Deane, Weichao Chen, Ngoc Vo
School of Information Science and Learning Technologies
Univeristy of Missouri-Columbia
kdgrz9@mizzou.edu, moorejoi@missouri.edu, cdickson-deane@mizzou.edu, wcxcf@mizzou.edu, ntv25c@mizzou.edu

Abstract: The purpose of this paper is to describe a designer's exploratory use of an e-learning usability evaluation heuristic in the beginning processes of designing and developing a nuclear math and theory online learning environment (OLE). A team consisting of a designer-developer and three evaluators provided feedback through an e-learning usability heuristic, expert background and experience, and/or visual design experience. Strengths and weaknesses of the process as well as the instrument are described and recommendations for the instrument and future heuristic evaluations are provided.

Introduction

Formative evaluation should start early and continue throughout the design and development of e-learning media (Gould, 1988). The evaluation process is often neither strictly formative or summative, but lies on a continuum between the two (Crowther, Keller & Waddoups, 2004). The earlier in the process of the evaluation, the more the evaluation will be formative, or provide feedback to iterate the current product; whereas evaluation later in the process will be more summative, or the product will be restricted to change (Crowther et al., 2004). Thus it becomes clear that the earlier on in the designing and development process one can intervene with usability guidelines and evaluation, the more information the designers and developers will have in regards to their design and the more impact it will have in regards to less-costly and time-intensive change in the design and increased usability.

Most of the literature focuses on two teams, each performing separate functions: one as designers or developers, and the other as the evaluation team (Bowman, Gabbard & Hix, 2002; Crowther et al., 2004; Federoff, 2002; Redish et al., 2002; Scholtz, 2004). This is understandable given the skills that are needed to perform each function; a designer is often not able to perform thorough usability evaluations and evaluators are not necessarily qualified to create designs or develop computer interfaces (Myers, 2004). However, what might be the experience and outcome if a designer were to design with an evaluator's heuristic checklist? How would having a heuristic checklist assist in the design and development process, and what might be needed?

The purpose of this paper is to describe the designer's exploratory use of an e-learning usability evaluation heuristic in the rapid prototyping design and development of a nuclear math and theory online learning environment (OLE). By examining the use of the heuristic in the early design and development, we can determine the strengths and weaknesses of the design and evaluation process and as well as the heuristic.

Background

The math and nuclear theory website (see Figure 1) is a problem-centered OLE intended for use in a web-assisted high school or undergraduate mathematics course (see Etter et al., in press). It was created to address a growing need for positions in the nuclear workforce in addition to enabling students to more greatly meet the rigors in nuclear mathematics courses. In high school settings, the course would be delivered over two semesters, or one semester for undergraduates, and therefore an entire year's curriculum and activities will be placed within the OLE.

The learning website contains a large amount of content that must be readily visible to the user and can be difficult to implement in regards to navigation, ease of use, and visual design (especially relating to "clutter"). Having a usability evaluation to assess how this content is organized, how the user is able to "keep track" of completed work, and the ability to move to different sections of the website, is of great interest in this formative usability evaluation.

The International Organization for Standardization (ISO) defines "usability" as "the extent to which the product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a

specified context of use," (Ergonomics, n.d.). Usability should be evaluated throughout the course of development rather than at the end through the use of usability engineering. "Usability engineering is the discipline that provides structured methods for achieving usability in user interface design during product development," (Scholtz, 2004). This paper addresses a small component in usability engineering: the use of heuristics in the formative usability evaluation of design.

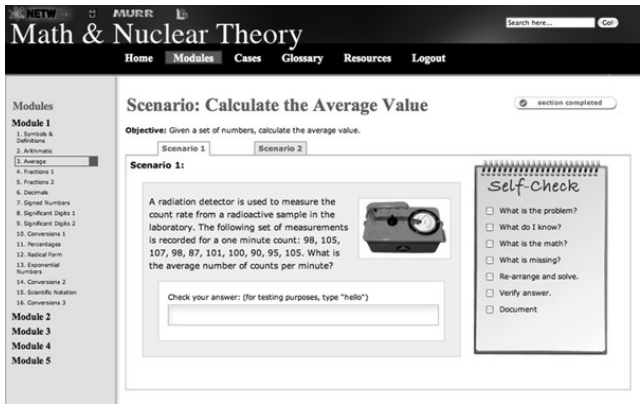

Figure 1. Sample page of the Math and Nuclear Theory Site

According to Scholtz (2004), formative evaluation is more about the design than taking actual measurements regarding usability, and is defined as user testing with the goal of learning about the design to improve its next iteration (Nielsen, 1994). The reporting is often informal, and performed in a way that expedites the process. Getting quick feedback in order to iterate the design and interaction is the main purpose of formative usability evaluation, and it drives the decision that must be made in the earlier conceptual and development stages (Rohn et al., 2002). Heuristic evaluation is one component of formative evaluation and is designed to be less time consuming and less expensive to conduct (Scholtz, 2004). While most evaluations only have trained evaluators utilizing a usability heuristic, in this case a designer, in addition to usability experts, will use the heuristic checklist.

It is stated that during the usability engineering process, the key is to test early and often (Henninger, 2000; Rohn et al., 2002). However, often projects are not funded to have teams available to support early-and-often testing, especially in-depth user-centered usability testing. One of the possible solutions is to have a small team of expert reviewers utilize a heuristic checklist in the beginning stages of the design.

There is much research regarding the use of heuristic checklists or guidelines in evaluation (Reeves et al., 2002; Mehlenbacher et al., 2005; Dringus & Cohen, 2005). However, most of them do not involve the designer utilizing a heuristic checklist themselves when creating the design (Nielsen, 1994). This could be that designers and developers are approaching the project from a different perspective of aesthetics and functionality, rather than utility. By understanding the different usability factors, the heuristics could be a meaningful guide for individuals with less knowledge and skills of an expert usability evaluator. However, by including usability as a component of design and development decisions, it could become difficult to manage several different aspects of the project by one person, rather than have different individuals focus on specific design and development elements. In essence, there may be times when there is an internal "conflict of interest" when designing the website. As such, we decided to focus on a designer's use of the instrument and have usability evaluators examine the website design with the same instrument.

Methods

The project team consisted of 4 members: a designer-developer, two usability evaluators, and one visual design evaluator. The designer-developer's role, as evident by the name, had both designing and developing roles in the creation of the OLE. For convenience, "the design process" will be referred to as the "designing and developing process", and "the designer" will be referred to as the "the designer-developer." The two usability evaluators, one is a professor specializing in learning technologies and the other, a graduate student specializing in usability in learning technologies, used the heuristic. The visual design evaluator was a PhD candidate in learning technologies with expertise in usability and visual design, herein referred to as the "visual design expert."

The instrument used for this formative evaluation was a heuristic checklist previously iterated by the team (see Moore, Dickson-Deane, Galyen, Vo & Charoentham, 2008). In this research, the fourth version of the revised

instrument included 12 categories and 38 questions. The evaluators first selected "yes" or "no" to indicate whether every heuristic principle was achieved and if not, they were instructed to provide comments and also indicate the severity and extensiveness level of the problem on a scale from 1 to 4 (low to high, respectively). At the end of the instrument, they provided a summary statement on the OLE as well as their perceived usefulness of the instrument in conveying usability issues.

Information regarding the usefulness of the usability heuristic in the design process was collected from both the designer as well as the evaluators. From the beginning of the design process, the designer had the usability heuristic available to refer to as needed over the course of approximately three months. In this situation, the designer only worked on the OLE for a few hours a week during a three-month time period. The designer was then interviewed regarding how and to what extent the usability heuristic was utilized during the design process as well as its perceived usefulness. After the initial design and functionality was created and viewable, the two usability evaluators conducted a heuristic evaluation on the learning environment using the same heuristic and the one visual design evaluator provided feedback. The usability evaluations and comments were analyzed in regards to content and usefulness in feedback regarding formative evaluations. The designer then was able to review the usability evaluations and visual designer feedback, and subsequently was interviewed regarding the perceived usefulness of the feedback, and how and to what extent the heuristic evaluations provided meaningful design recommendations.

Analysis

Analysis of the data fell into three main categories: the designer's use of the instrument, the evaluators' use of the instrument, and the designer's perceived usefulness of the evaluators' feedback (presented in the form of the instrument). The designer used the instrument from the beginning of the process and indicated that it was a useful tool. The main benefit was that it provided a checklist of usability guidelines while including website features from the beginning. However, the designer felt that the usability heuristic did not have the ability to address all major aspects of creating a learning website.

> *"I think the heuristic is a good tool for addressing the science of creating a usable site, but not so much the art of creating a usable site. There is a lot of science to it, but there's also the art of it that impacts the satisfaction and affective impact a user has when interacting with the site."* [D1]

Likewise, the evaluators commented on similar aspects of the usability heuristic to address aspects of visual design, which may be difficult to capture in guidelines:

> *"Overall, I felt like I gave a 'surface' response, not really going into detail. I think the design looks fine, but would it be important to rate it on a likert scale? There is such a big range between yes and no."* [UE1]

> *"The lack of emphasis on visual design of this evaluation instrument may have influence on the design. This may explain why this website is not strong in visual design."* [UE2]

The evaluators seemed to feel constrained by the heuristics, because it did not address some of their concerns regarding visual design. One of the evaluators cautioned that visual design can be highly subjective, as was evident in the large difference between the two usability evaluators opinions regarding visual design issues. One also noted that visual design as well as some of the items in the instrument can be highly context-dependent. This is supported by Henninger (2000) in their investigation of context-specific usability guidelines. The feedback received from the visual designer, however, rated the visual design high, but it was not in-depth and was cursory compared to the feedback given by the usability evaluators using the heuristic.

The feedback from the usability evaluators and visual design evaluator was given to the designer to help in redesign. However, the designer felt that the heuristic feedback was reiterating what was known, often presented with problem lists rather than redesign suggestions. As a "one-man band" designer-developer, the designer preferred more problem identifications with possible solutions for redesign. In addition, some problems or comments were interpreted as vague or confusing. For example:

> *"...Some of the suggestions made me wonder what they were seeing. For example, images had clear backgrounds but the evaluator continually stated that the images were different colors from the background, causing it to look visually unpleasant."*[D1]

"Both of the evaluators stated that when you logged out and logged back in, it didn't bring you back to the page where you were last at. It works on everything I've tried. I am curious what they experienced; I'd like more feedback on what they were using or seeing. I don't know how to recreate what the evaluators were seeing." [D1]

The designer received feedback from the visual designer, which lacked the in-depth usability checks as well as the constraints of the heuristic. While the designer found the visual designers' feedback more cursory and less detailed, it was felt it did have some of the "artistic" feedback that the heuristic lacked.

"[The visual designer feedback] was not as extensive as what the heuristic provided, but I feel like it addressed some of the "art" of usability. She told me what she liked, what she felt could be better, what might cause confusion, and some suggestions for changing. I especially liked the suggestions for change. Even if I don't go 100% with what the person suggests, the suggestions give me a clearer idea of what she was thinking was inappropriate, wrong, or confusing with the interface and helps me to think along different lines or even spurs me on to a better idea than the suggestion." [D1]

Discussion

There is much information on how to conduct evaluations but not on reporting, especially heuristic reporting (Theofanos & Quesenbery, 2005). How to present the findings from the heuristic can make the difference between a highly useful or meaningless evaluation. Horbaek & Frøkjær (2005) described the difference between problem lists and redesign proposals. They stated that in much of the literature, a good usability instrument is assisting the evaluator in the generation of a problem description, rather than in using the problem description to facilitate some potential redesign solutions. In their mixed-methods study of developers receiving problem lists versus suggested redesign proposals, developers valued redesign proposals over problem lists as input to their work. They also found that for much of the data, the usability evaluations, when presented as problem lists, were often confirmations of what they already knew.

This seemed to occur during this case with the use of the usability heuristic. While the use of an e-learning usability heuristic in the beginning stages of design and development was beneficial in helping the designer implement important usability features as early as possible, it may have rendered the expert formative usability evaluations as having less impact. As Hornbæk and Frøkjær (2005) noted in their work, the problem lists resulting from the heuristic evaluation were familiar to the designer. For greater usefulness and impact, they suggested it may be more useful to supplement problem list with redesign suggestions. This benefits the designer both in receiving more information from the heuristic evaluation through redesign suggestions as well as reducing vagueness of problem issues by providing sample solutions.

An alternative to this solution is to consider the heuristic evaluation as a "part 1" of the heuristic evaluation, with "part 2" being the meeting of the team to discuss results. In this manner, the designer can ask about comments that seem vague or confusing, attempt to recreate problems specified, and brainstorm about potential solutions.

Another method would be to use a Delphi evaluation method where the process of evaluation will be done individually with an evaluation lead to summarize and re-introduce the summary to the individual experts/designer. This way consensus can be gained through multiple evaluation attempts, and it reduces the level of ambiguity between evaluative reports (Wysocki, 2003).

The usability experts and the designer stated that certain visual design aspects were difficult to assess with the heuristic. These results also seem to support that many usability instruments, especially heuristics, have difficulty addressing the affective domain, or "satisfaction," (Zaharias, 2004; Zaharias, 2006). By including a visual designer on the team, unrestricted by a heuristic, seemed to address some of these issues, though more cursory than what would have been preferred. A solution to this void in the heuristic could be to incorporate visual designer feedback with heuristic results in order to help capture the artistic and affective component that the heuristic seems to lack.

Acknowledgements

The Nuclear Math and Theory online learning environment is part of a one-year grant. The team is composed of the following hardworking individuals: Randy Etter, William H. Miller, Gayla Neumeyer, and Lori Kilfoil from the University of Missouri Research Reactor; and Bruce A. Meffert, Theresa J. Yeager, and Victoria A. Schwinke from Linn State Technical College.

References

Bowman, D. A., Gabbard, J. L., & Hix, D. (2002). A survey of usability evaluation in virtual environments: Classification and comparison of methods. *Presence: Teleoperators & Virtual Environments*, *11*(4), 404-424.

Crowther, M. S., Keller, C. C., & Waddoups, G. L. (2004). Improving the quality and effectiveness of computer-mediated instruction through usability evaluations. *British Journal of Educational Technology*, *35*(3), 289-303.

Dringus, L. P., & Cohen, M. S. (2005). An adaptable usability heuristic checklist for online courses. In *Proceedings of ASEE/IEEE Frontiers in Education Conference*, Indianapolis, Indiana, 19-22 October 2005 (pp. T2H-6 – T2H-11). Piscataway, NJ: IEEE.

International Organization for Standardization. (n.d.). *Ergonomics of human-system interaction -- Usability methods supporting human-centred design.* Retrieved April 30, 2009 from ISO Web Site: http://www.iso.org.

Etter, R., Miller, W., Neumeyer, G., Meffert, B., Yeager, T. J., Schwinke, V. A., Galyen, K. (in press). Development of a problem-based undergraduate nuclear course to strengthen math skills in support of the US nuclear workforce. *Transactions - American Nuclear Society.*

Federoff, M. A. (2002). *Heuristics and usability guidelines for the creation and evaluation of fun in video games.* Unpublished Thesis.

Gould, J. (1988). How to design usable systems. In M. Helander (Ed.), *Handbook of human-computer interaction*. North Holland: Elsevier Science.

Henninger, S. (2000). A methodology and tools for applying context-specific usability guidelines to interface design. *Interacting with Computers*, *12*(3), 225-243.

Hornbæk, K., & Frøkjær, E. (2005). Comparing usability problems and redesign proposals as input to practical systems development. In *Proceedings of the SIGCHI conference on human factors in computing systems.*

Mehlenbacher, B., Bennett, L., Bird, T., Ivey, M., Lucas, J., Morton, J. & Whitman, L. (2005). Usable E-learning: A conceptual model for evaluation and design. In *Proceedings of HCI International 2005: 11th International Conference on Human-Computer Interaction*, Las Vegas, Nevada, 22-27 July 2005.

Moore, J.L., Dickson-Deane, C., Galyen, K., Vo, N. & Charoentham, M. (2008). e-Learning Usability Instruments: What is being Evaluated?. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2008*, Las Vegas, Nevada, 17-21 November 2008. (pp. 430-435). Chesapeake, VA: AACE.

Myers, G. J. (2004). *The Art of Software Testing.* New York: John Wiley and Sons.

Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods* (pp. 25-62). New York: John Wiley & Sons.

Redish, J. G., Bias, R. G., Bailey, R., Molich, R., Dumas, J., & Spool, J. M. (2002). Usability in practice: Formative usability evaluations-evolution and revolution. In *Conference on human factors in computing systems.*

Reeves, T. , Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E. & Loh, S. (2002). Usability and instructional design heuristics for e-learning evaluation. Paper presented at *ED-MEDIA 2002, World Conference on Educational Multimedia, Hypermedia & Telecommunications.* Denver, Colorado. Retrieved April 29, 2009 from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/19/c8.pdf

Rohn, J. A., Spool, J., Ektare, M., Koyani, S., Muller, M., & Redish, J. G. (2002). Usability in practice: Alternatives to formative evaluations-evolution and revolution. In *Conference on human factors in computing systems.*

Scholtz, J. (2004). Usability evaluation. *Retrieved October*, *13*, 2005.

Theofanos, M., & Quesenbery, W. (2005). Towards the design of effective formative test reports. *Journal of Usability Studies*, *1*(1), 27-45.

Wysocki,R. (2003). *Effective Project Management: Traditional, Adaptive, Extreme.* New York: John Wiley & Sons.

Zaharias, P. (2004). Developing a usability evaluation method for e-learning applications: From functional usability to motivation to learn. *Retrieved September*, *8*, 2007.

Zaharias, P. (2006). A usability evaluation method for e-learning: Focus on motivation to learn. In *Conference on human factors in computing systems.*