Reference: Proceedings of the ED-MEDIA 2005 annual conference, Montreal, pp. 1162-1167 (2005, June).

Improving the Design of Web-based Games and Simulations through Usability Research

Ron Owston
rowston@edu.yorku.ca
York University Institute for Research in Learning Technologies

Andre Kushniruk
andrekushniruk@usa.net
University of Victoria, BC

Francis Ho
francho@shaw.ca
York University Institute for Research in Learning Technologies

Kevin Pitts
Kevin_pitts@senecac.on.ca
Seneca College of Applied Arts and Technology

Herb Wideman
herb@yorku.ca
York University Institute for Research in Learning Technologies,

The Virtual Usability Laboratory (VUL) is a software tool developed to remotely capture and analyze a wide variety of usage data on Web-based educational games and simulations. The VUL allows for automated remote collection and integration of such data as user activity logs, on-line demographic questionnaire responses, and data from automatically triggered pop up questions at critical points in the software application. Collected data are stored in an integrated database system, allowing for subsequent data mining and ad-hoc querying by researchers. We illustrate the utility of the VUL in providing feedback to the designers of the educational game Trivia.

Introduction

A recent review of research conducted by the Ultralab and the Learning and Skills Development Agency in the UK (Mitchell & Savill-Smith, 2004) pointed out the critical need for more rigorous research methodologies to usability and outcomes of educational games. In the Simulation and Advanced Gaming Environments for Learning (SAGE) project (http://sageforlearning.ca), we are focusing on the usability component of this problem. Our goal is to develop a software application to collect and store for subsequent analysis, user activity and perceptions when playing a game. The purpose of this paper is to report on our progress in the development of this tool and to illustrate how it can be used to provide feedback to game developers on learner perceptions and usage patterns of their applications.

Well-designed educational games and simulations appear to have significant potential for heightening student engagement in learning and enhancing a range of learning outcomes. The latest generation of computer technologies are now capable of supporting networked, collaborative simulation and gaming environments that enable the situated embodiment of the major learning principles found in empirically-supported cognitive and social learning theories. Recent applications have ranged from the custom development of a 3D immersive simulation for business leadership training (Aldrich, 2004) to the repurposing of a commercial strategy game, Civilization III, to the teaching of American history (Squire, in press). A growing body of research supports the proposition that by providing for interactivity, immersion, collaboration, learner control, intrinsic motivation, and learner reflection and problem solving, these new environments can enhance cognitive and social learning (see Kirriemuir & McFarlane (2004) for a recent literature review). Some researchers suggest that educational gaming and simulation environments are ideally suited to meet the learning needs of a younger generation characterized by a greater capacity for multitasking, a shorter attention span, and greater experience in learning via multimedia enriched exploration and discovery (e.g., Gee, 2003; Shaffer et al 2004). Learning through games and simulations has the potential to transform key elements of traditional classroom pedagogy, as students experiment or play first, and then inductively develop deeper understandings and theoretical knowledge based upon the consequences of their actions

(Saiethang & Kee 1998). Teacher and learner roles can be shifted, as students become more active and assume a collaborative teaching role with their peers.

To realize these potentials in any given game or simulation, developers must pay careful attention to issues of design and usability. The latter is especially critical in educational applications, so as to reduce the cognitive imposition of engaging the software's structure and interface on the learning process. The goal here is to minimize the cognitive load on learners, and in so doing leave more cognitive resources free to focus on substantive learning (Bransford, Brown, & Cocking, 2002). Understanding the processes by which a learner interacts with both content and design of the game or simulation, and using that knowledge to improve the efficacy of the software, requires ongoing study of potential end users in an iterative development cycle incorporating needs assessment, design, prototyping, usability assessment, and formative evaluation (Kushniruk & Patel, 2005). The emergent understandings of users' cognitive processes, including their substantive misconceptions as well as any difficulties encountered due to delivery and design factors, can be used to guide subsequent revisions of content, work processes, and user interfaces. An important methodology for analyzing situated cognition and decision making known as *cognitive task analysis* (Gordon & Gill, 1997) has been successfully applied to study the reasoning and decision making of computer-based knowledge systems. In cognitive task analysis a number of techniques are used to document the thinking of users *in situ*, including capturing all of the users' computer screen actions to video, along with a synchronous audio recording of the users as they "think aloud" while working. While providing extremely rich data for understanding the work processes of end users, analyzing this data is very time-consuming and expensive, limiting the number of subjects that it can be used with. And because no prompts or questions are given users, they may not offer any reflections or opinions on topics of critical interest to developers (for example, "How easy was it to navigate the cascading menu system? Were the headings and subheadings logically organized? Did the animated graphic you just viewed provide you with a clear understanding of the hydraulic motion in the pump?") This paper presents a second approach to user task analysis that can either supplement or partially substitute for traditional "think-aloud" protocols—that incorporate the use of remotely delivered, automated online queries customized to capture critical user experiences and perceptions at key points in the use of gaming and simulation environments.

The approach is embodied in a tool we are developing, known as the Virtual Usability Laboratory (VUL). The tool has been modified and extended for the evaluation of Web-based educational games and simulations from its original use in assessing health care information systems (see Kushniruk & Ho, 2004). The VUL allows for the automated integration of a wide range of sources of data, ranging from user activity logs, online demographic questionnaire responses, and data from automatically triggered pop-up questions, to the results of queries designed to automatically appear at key points when users interact with games and simulations. Another feature of the tool is its capability to record the screens and voice conversations of remote users and store the files on the VUL server. The data that are collected are stored in an integrated database system, allowing for subsequent data mining and ad-hoc querying of the data by researchers. The VUL also allows for ease of use for researchers in setting up the parameters for studies and automatically monitoring users whether they are playing games locally or are scattered across the Internet.

Design of the Virtual Usability Laboratory

Our approach to developing the VUL has involved the creation and integration of five interacting system components:

1. **Central tracking component**: This component resides on an evaluation server (a computer located in our facilities), for remotely tracking and logging use of Web-based games and simulations located at remote sites. This component, programmed using ColdFusion, can provide a customized (i.e., researcher defined) log file of accesses by users to a system under study. It can provide a log of what Web pages within a site are accessed, the order of browsing, and a time-stamped record of the users' activities in accessing a remote site. For example, using our tracking system (physically located at our facilities in city X) we can track a user (located in city Y) accessing a site being evaluated which is physically located at a third site (city Z). All that is required is providing the tracking component with the Web address (URL) of the remote site which is being evaluated. All subsequent user accesses to the target site are then transparently redirected through our evaluation server, i.e., the VUL tracks users interacting with a Web site by rerouting their interactions through the tracking component of the evaluation server (assuming user informed consent is obtained). This approach to logging use of a remote Web site has the following advantages: (1) a

customized log file of a user's interactions with a Web site can be created without having to alter or create a log file at the site being evaluated; and (2) by rerouting user requests through the evaluation server, the ability to query users about their access to sites and pages they are about to go to, or they have just returned from, becomes feasible, as described below.

2. **Screen and voice capture**: In fast action games and simulations and those with modules designed in Flash or Java, remote tracking of all user interactions via logging is not practicable. Therefore, we have a component that records the screens of remote users for the entire user session. Currently, this component uses Macromedia Breeze Live (http://macromedia.com). For team playing and when users are asked to talk aloud about their playing, audio can be recorded as well. As Breeze does not allow these recordings to be exported, the recordings are played back by the researcher and captured with local screen capture software. The files are then tagged and stored in the VUL database (see 4 below) and can be analyzed by the researcher using video analysis software such as Transana (http://www2.wcer.wisc.edu/Transana/Transana).

3. **Database component**: This component allows for storing and integrating the results of online logging and the results from presentation of online questionnaires in a relational database containing other information about users including demographic and educational background information, using relational database tables to store, interrelate, and query usage data. To accomplish this, data obtained from user interactions are stored in separate database tables. For example, a table can be created containing the results of logging of users' interactions with a web site (with the fields of the table consisting of attributes such as time entered the site, time left the site, URL, and user ID). This type of information can be relationally linked with other tables that are also automatically generated containing information about user demographics, results of on-line questionnaires, etc. By linking across key fields such as user ID and time, a wide variety of ad-hoc querying (using basic SQL commands) is allowed for in the researchers' examination of the evaluation data stored in the tables. In addition, the creation of linked data tables allows for the possibility for application of data mining and data discovery algorithms for identifying emergent patterns from integrated usage data.

4. **Presentation component**: This component is for controlling the presentation of sequenced on-line questionnaire forms for users to complete to assess the usability of Web sites remotely at point of use. The triggering of such questionnaires can be based on a user profile created for each user of a site being evaluated. By redirecting requests for access to a Web site under evaluation through our evaluation server (as described above), we are able to write scripts that can trigger questionnaires and prompts for user information to occur at points when users enter or leave parts of a site of interest. For example, on first entry into a system under study we can pop up a demographic questionnaire and later, when the user accesses a page of interest (e.g., a page containing additional information about how a simulation works), an on-line questionnaire can also be made to appear to query the user about why a page is accessed, satisfaction with information provided, educational value of the content provided, etc. In addition, questionnaires and multiple choice tests of knowledge can be triggered to appear after a certain number of invocations of an educational program or at key points in user interactions with such applications, as specified by the researcher (described below).

5. **Researcher interface component**: A user interface component was designed that allows evaluators of a game or simulation Web site to easily set the parameters of the evaluation, without knowing the implementation details of the VUL. For example, a researcher may specify what site will be evaluated (by providing the site's URL), what type of questions should appear to users (by adapting or editing questionnaires contained in a questionnaire bank), and indicating when questionnaires/logging is to be triggered. The process of creating a Web-based evaluation involves the researcher accessing the researcher's component of the VUL. The researcher is then prompted to enter the URL of the site to be evaluated and is prompted to enter the format of any required log files to be created to record user interactions with the site under study. The researcher can then indicate what type of initial background questionnaires or prompts should appear when users first enter the site (if any), as well as to indicate if any particular prompts or questionnaires should appear to users as they enter specific pages or functions within the site. To facilitate this process, the researcher can choose from and modify questionnaires and prompts contained in a template bank, or alternatively choose to create their own. Finally, the researcher is

prompted to indicate what type of data analysis they would like by selecting from a list of built-in types of statistical analyses. At this point the parameters of a specific remotely conducted Web-based evaluation are set and the data collection from users of the site can proceed. The actual rerouting of the users' interactions with a system through the VUL can be achieved by having users enter a URL which will transparently redirect users through the VUL when they interact with the Web site under evaluation.

The architecture of the system is illustrated in Figure 1 below.

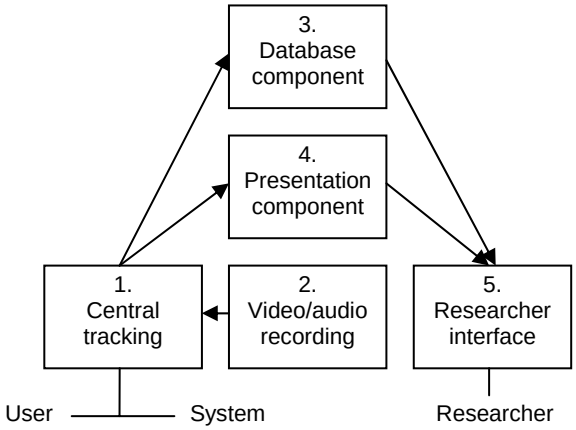

Figure 1: Overall System Architecture

Application of VUL to Study Games and Simulations

The VUL is currently being tested and deployed on several projects examining different kinds of educational games and simulations. One illustration is our use of the VUL to remotely record user interactions at the Educational Games Central website, which allows teachers and trainers, companies and community or non-profit organizations to develop and use educational games online using generic game shells (see http://www.savie.qc.ca/carrefourjeux/an/accueil.htm). Five shells based on popular board games are available at this site. We have been testing the VUL with the Trivia shell, which requires the instructor to write a minimum of 54 questions—nine questions in each of six categories. From two to six players can participate in a game. Our field trials have been conducted in a computer lab session of a college level introductory information technology management course. The game is used to review particular course topics covered in lectures.

We have been using the Breeze Live recording component for these trials (see component 2 above). Since Breeze cannot be launched by external routines, users first log into Breeze. They are then provided with a sequence of steps to follow, beginning with turning on the Breeze screen recording feature and clicking on the url for the VUL. Pre-game questions pop up asking for demographic data and computer familiarity. (These questions can be readily modified by the researcher.) Next users log into Education Games Central, select the Trivia game, and play it. At the end of the session they log out and click on the link to the VUL. VUL, in turn, prompts them with post-game questions. After they complete these, users turn off the screen recording and log out. A research assistant is available on site to provide help in following all steps of the trial, however our goal is to automate most of them so that on site assistance will not be needed.

Once the recordings are made, the researcher is responsible for capturing them with screen recording software and entering them into the relation database described above in component 4. (Our goal is to automate this step as well.) This allows the research to query the database on specific user attributes (e.g., level of computer experience) and retrieve the videos of all users with, for example, minimal computer experience. Another query might be on games scores, which would let the researcher view, say, videos of high and low scorers to compare strategies that they employ.

Below are two screenshots taken from a recording of a user session to provide an illustration of what the player sees.

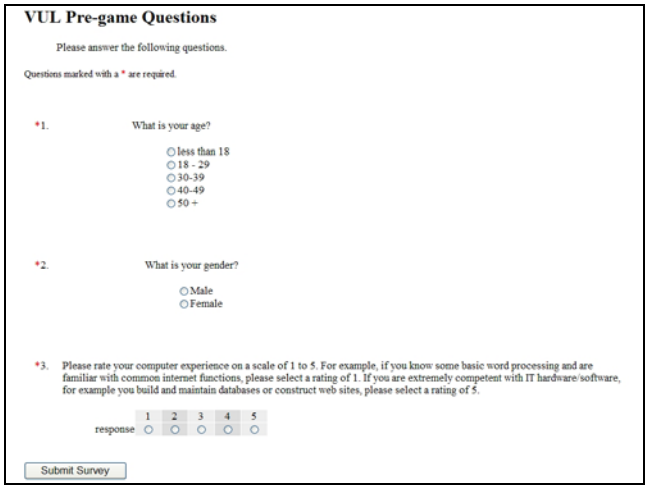

Figure 2: VUL Pop-up Pre-game Questions

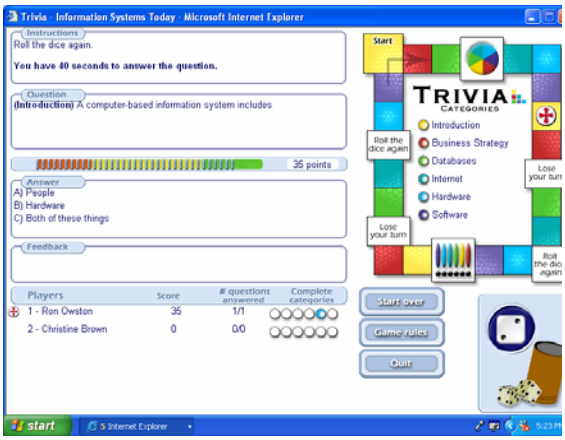

Figure 3: Sample Question from Trivia

Overall, we have found that the collection of varied forms of usage data remotely is not only feasible, but additionally by storing data in a consistent format (e.g., a relational database system) the integration and querying of varied forms of usage data can be supported for practical purposes. By using the VUL we are able to provide iterative feedback to the developers on the use and usability of components of their application, assess user perceptions about features of the software, and identify and assess key aspects of games and simulations that could be incorporated into other applications of this type.

Conclusions and Further Research

In conclusion, we believe that the results of our work on the VUL will be significant for educational researchers who are charged with assessing not only simulations and games, but any kind of interactive Web site as it will provide them with a relatively simple to use tool than can be customized to collect and analyze a wide variety of user data. Our next development steps include the creation of a video capture routine that does not rely on Breeze, load testing of our server to determine the number of simultaneous users that can be tracked, provision of a database of usability questions for researchers to draw up when setting up their own studies, and improving the interface of the researcher component of the system[1].

References

Aldrich, C. (2004). *Simulations and the future of learning.* San Francisco: Pfeiffer.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (2002). *How people learn: Brain, mind, experience, and school*. (expanded ed.) Washington, D.C.: National Academy Press.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy.* New York: Palgrave MacMillan.

Gordon, S.E., & Gill, R.T. (1997). Cognitive task analysis. In: Zsambok, C.E., Klein, G., editors. Naturalistic Decision Making. Mahwah, NJ:Erlbaum, 131-140.

Kirriemur, J., & Mcfarlane, C.A. (2004). *Literature review in games and learning: A report for Nesta Futurelab.* Retrieved November 5 2004 from http://www.nestafuturelab.org/research/reviews/08_01.htm.

[1] More information on the development of the Virtual Usability Lab can be found at http://www.vulab.ca/.

Kushniruk, A. W., & Ho, F. (2004, May). *The Virtual Usability Laboratory: Evaluating Web-based health systems.* In Proceeding of e-Health 2004, Victoria, BC.

Kushniruk, A.W., & Patel, V.L. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics, 37*, 56-76.

Mitchell, A., & Savill-Smith, C. (2004). *The use of computer and video games for learning: A review of the literature*. London: Learning and Skills Development Agency. Retrieved April 28, 2005 from http://www.lsda.org.uk/.

Saethang, T., & Kee, C.C. (1998). A gaming strategy for teaching the use of critical cardiovascular drugs. *Journal of Continuing Education in Nursing*, *29*(2), pp. 61-65.

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2004). *Video games and the future of learning.* University of Wisconsin-Madison and Academic Advanced Distributed Learning Co- Laboratory: December. Retrieved March 30, 2005 from http://www.academiccolab.org/resources/gappspaper1.pdf.

Squire, K.D. (in press). Civilization III as a world history sandbox. To appear in *Civilization and its discontent: Virtual history, real fantasies*. Milan, Italy: Ludilogica Press.