Is it Ethical to Evaluate Web-based Learning Tools using Students?

M.-A.D. Storey[1] B. Phillips M. Maczewski
Departments of Computer Science and Psychology
University of Victoria, British Columbia, Canada

In the Spring semester of 2000 we conducted a controlled and situated study evaluating web-based learning tools. While conducting this study, ethical issues concerning the relationships among research participants (students) and experimenters (instructors and graduate students), research process (informed consent, minimization of harm, competence and confidentiality) and unevaluated tool deployment within university class settings arose. In this short discussion paper we will describe several of these situations and some of the ethical questions that arose from them.

Background

Web-based learning tools are experiencing widespread adoption across North America at university and college campuses. Learning tool advocates suggest that the tools will improve the quality of education for students, satisfy diverse learning skills, support instructors unfamiliar with web technologies, and ensure consistency across departments and universities. However, the potential disadvantages of applying these tools are significant. Some educators are concerned that these tools are being deployed without regard for students' and instructors' well being. Meanwhile, institutions are adopting these technologies without fully understanding their impact on the institution, administrators, instructors, and students.

To answer some of these questions, our research group conducted a controlled study of these tools in a university classroom. We wanted to know if the students, instructors, and course administrators found the tools difficult to use and whether they should be deployed across campus. Moreover, we were interested in the students' perceptions of how these tools impacted their learning.

We evaluated and compared two tools, WebCT [1] and Blackboard [2], in a third year course on Human-Computer Interaction at the University of Victoria. 54 (out of 57) students volunteered to participate in the study and were randomly assigned to two groups of 27. A simple web site (constructed by the instructor) was used for the first four weeks of term. This website provided access to course announcements and course material, but had no special features common to web-based learning tools such as for submitting assignments, quizzes, or communication tools. After the first four-week period in the term each group used both Blackboard and WebCT, in counterbalanced order, for four weeks.

Prior to this study we obtained permission from the Human Research Ethics Committee at the University of Victoria to use students in the study. Despite approval and our efforts to make a comprehensive ethical review of our research, there were still some minor concerns that arose over the course of the project. Vinson and Singer in their paper on Ethical Issues in Empirical Studies of Software Engineering [3] describe four common issues found across different codes of ethics: *informed consent, minimization of harm in relation to merit, competence and confidentiality*. In our discussion of our ethical concerns we will limit ourselves to these four issues.

[1] email mstorey@uvic.ca.

Informed Consent

Free and informed consent refers to the dialogue, information sharing and general process through which prospective subjects choose to participate in research. At the beginning of our study we explained to the students the procedure, equipment, and possible benefits and disadvantages of participating in the study. We feel students were completely informed about the study, but are not sure whether the consent we obtained was completely free.

We may have inadvertently coerced students into participating in the study in a couple of ways. Students could gain course credit by participating. Part of the rationale for this was that they would gain experience evaluating user interfaces and would gain the experience of being a participant in a research study. There was an alternative assignment (to write a review of the interfaces of the two tools) for those who decided to not participate. This option may have been less attractive to the students (computer science students are notorious at hating to write essays). Perhaps this was coercion? We did, however, try to make the alternative assignment as time consuming and potentially as academically valuable as participating in the study.

Another criterion for the study was that students could withdraw at any time, without penalty. Our study may have, in fact, created the opposite scenario, making dropping from the study more attractive then participating in it. As one student mentioned, those students opting to withdraw from the study did not have to complete the assignment, and this was unfair to those selecting to opt out from the beginning. How to handle this was not entirely clear.

Also, the instructor and the teaching assistant were both involved as experimenters. This could possibly have put undue stress on the students to participate (they may have perceived that non-participation could affect their grade). However, we do not believe it would have been possible to conduct such a study without this tight connection between instructors and experimenters. We were intimately knowledgeable about the course material and deliverables and this affected how we kept the tools up to date. Perhaps it would be possible to construct a study without this tight integration, but the results may not be as strong and the potential harm (see next section) may actually increase.

Minimization of Harm

Compared to non-participants, those participating in the study had the additional stress of learning two new web-based learning tools and switching between them. Would this impact their learning process? On the one hand, there was the advantage gained of learning and critiquing different user interfaces (the focus of the course) and being a subject in a user study (one of the deliverables of the course was to design a user experiment, apply for ethical review and consequently implement the experiment). On the other hand, the students that did not participate in the study were required to critique the user interfaces and write a report. They did not gain the experience of what it is like to be a subject in a user study.

As we were maintaining three different environments, the researchers spent an immense amount of time making sure that all three tools were running and accurate. Also, the tools had some bugs. Consequently, we had someone available to make updates to the site and respond to student queries almost 20 hours a day. We would like to note that normally instructors are not available this much and sometimes web pages can have broken links or servers can be down. What are the researchers responsibilities in regards to technical support when conducting research with software tools within a 'live' class?

Although we did not directly measure the effect of these tools on learning (we did everything possible to mitigate any problems to avoid this sort of harm), some students felt that the tools did affect their ability to learn. So does this mean that when these tools are deployed they could possibly cause harm? The version of WebCT we evaluated did not fare well in our study. WebCT is the standard web-based learning tool at the University of Victoria. Following our study, the university has struck a task force to investigate the affects of using this tool.

Competence

Although most of the experimenters had computer science backgrounds, we were not intimately familiar with the tools deployed. Even though we had a strong technical background, we found the tools hard to use, especially in the administrator role. We were also inexperienced in setting up the tools for course use, so perhaps did not choose the best options. This could have biased the evaluation (and hence our results) and could have potentially increased the risk of harm to students. This was something we did not consider before the study. These tools are geared at instructors that have even less expertise than we did. What are the criteria for assessing 'competence' when software tools are used within a study? Is the 'competence' (usability) of a web-based learning tool part of this assessment?

Confidentiality

All of the data from the study was collected and stored electronically. Although the submission of electronic questionnaire data was anonymous, it would have been possible to identify where the questionnaires originated using server logs of IP addresses. Obviously, we did not do this, but the potential is there.

Another issue concerning confidentiality was that the web based learning systems tracked students' use of the tools. At the beginning of the study we were not aware that the tools did this tracking. In fact, the promotional material we saw beforehand did not mention this feature (although later when the course was in progress this feature was discussed at a public demonstration of the tools as being available and was suggested that it could be used to monitor students not doing well!). We were appalled that this information was being collected and did not access it at any point during the study. We did use aggregate data (which was also available) to distinguish and compare trends between the tools. Were we being unethical evaluating these tools before fully understanding all the functionality these tools had to offer? Moreover, is it ethical that these tools track this kind of personal information without students' knowledge? Couldn't an instructor perhaps unethically use such information to help decide a student's grade?

Conclusions

Pragmatically, it is important to recognize that there are many obstacles and ethical dilemmas to conducting these kinds of situated usability studies. Consequently, data and feedback from classroom-based users is difficult to collect. Also, evaluating distributed learning tools under development in a classroom setting has ethical implications that have to be considered by the researchers. In an effort to reduce the potential harm (e.g., lower class grades or stress) to students using untested software, we evaluated the tools using computer science students with a background in user interface design. The results from our study may have been different if we had drawn our student participants from a different discipline. We purposely chose this student group as they were familiar with a wide variety of computer applications and hopefully would not be unduly burdened by using these tools. However, the fact that they experienced difficulties further increases the likelihood that students from a non-technical domain

would also have experienced difficulties. Although some of the ethics of our study may be questionable, if there is a potential that these tools would cause stress and affect learning negatively then how ethical is it to deploy these tools within a university environment?

References

1. WebCT Homepage, http://about.webct.com .
2. Blackboard CourseInfo Homepage http://company.blackboard.com.
3. Ethical Issues in Empirical Studies of Software Engineering by Vinson and Singer.
4. Evaluating the Usability of Web-based Learning Tools, by M.-A. Storey, B. Phillips, M. Maczewski, and M. Wang. Accepted for publication at Ed-Media 2001, World Conferences on Educational Multimedia , Hypermedia & Telecommunications, Tampere, Finland, June 2001.