

# Is it Ethical to Log Users' Actions in Software Engineering Experiments?

*Amela Karahasanovic*  
*Simula Research Laboratory, Lysaker, Norway*

[amela@simula.no](mailto:amela@simula.no)

## Abstract

In recent years, there has been a significant increase in the number of empirical studies in software engineering that involve human participants. Such studies have heightened the need for considering ethical issues. Researchers have not only moral, but also pragmatic, reasons for treating participants in their studies ethically. They want to maintain their access to the data source and to get funding for their research. Thus, the researchers need to be aware of both ethical issues in general, and those that are specific to a particular research practice.

This paper briefly describes our experiences with a tool that logs all actions of the users/participants (commands, keystrokes and mouse movements), and collects their comments on a web-based screen. The tool was used in two usability studies and raised ethical issues concerning the assurance of confidentiality and minimization of inconvenience. The paper also discusses a conflict that appeared during these studies, between ethical standards and the ensuring of experimental validity.

**Keywords** : ethics, experiment, empirical software engineering, logging

## Introduction

The term 'ethics' denotes here the standard for moral behavior of a particular group, such as "the ethics of physicians" or "the ethics of engineers" (Onlineethics.org, 2001). Considering ethical issues is an important part of the planning and conducting of empirical studies in software engineering, for moral and pragmatic reasons (Singer & Vinson, 2001). To behave ethically, researchers need to be familiar with existing standards for conducting responsible research (Sieber, 1992) and standards for practicing software engineering, such as the IEEE-CS/ACM "Software Engineering Code of Ethics and Professional Practice" (1999). As pointed out by Singer and Vinson (2001), the researchers must also identify and address ethical issues that are specific to a new research practice.

We have developed a tool that automatically collects information about the interaction between the participants of a study and a computer system (Karahasanovic, Sjøberg, & Jørgensen, 2001). The following operations are logged, together with timestamps: windows operations, keystrokes, operations on the mouse buttons and *unix* commands. This tool also presents to the participants a web-based screen for entering data at regular intervals. The participants write down their thoughts and experiences on the screen, and the tool collects their feedback, together with timestamps. This tool was used in two usability studies (Karahasanovic & Sjøberg, 2001, 2002). The first study was a controlled experiment with 13 students and the second was a long-term exploratory study with one professional. The tool facilitated data collection and analysis, but at the same time raised ethical dilemmas. What new kinds of risk might the use of this tool create for participants in the experiment? How might it be

---

Material published as part of these proceedings, either on-line or in print, is copyrighted by Informing Science. Permission to make digital or paper copy of part or all of these works for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage AND that copies 1) bear this notice in full and 2) give the full citation on the first page. It is permissible to abstract these works so long as credit is given. To copy in all other cases or to republish or to post on a server or to redistribute to lists requires specific permission from the publisher at [Publisher@InformingScience.org](mailto:Publisher@InformingScience.org)

possible to deal with these risks? Are there any conflicts between our interests as researchers and the interests of participants in the experiment? If so, what are they, and how might they be dealt with?

## Risks for the Experiment Participants

Sieber has identified several risks for participants involved in empirical studies in software engineering: inconvenience, psychological, social, economic, legal and physical (2001). Inconvenience is constituted by, for example, boredom and frustration. Psychological risk is, for example, to worry about breaches of confidentiality and evaluation of the conducted work. Social risk may be disapproval by colleagues or superiors and economic risk may be loss of employment. Legal risk is, for example, a lawsuit, and physical risk is, for example, injury. We find the first three most relevant for our studies and will limit our discussion to them.

- *Inconvenience*: The participants of our studies used a research prototype. The prototype lacked proper documentation and a help function. This might be frustrating for the participants as users. The logging tool interrupted the participants in their work. This might be irritating. Further, the participants might feel that they were wasting their time during the studies. The students in the first study might regard their participation as taking valuable time away from their studies. The professional in the second study might regard his participation as preventing him from spending time with his family, because he participated in the study after working hours.
- *Psychological risk*: The participants might feel that they were not only observed, but also evaluated. The students might worry that the evaluation would affect their grades and employment opportunities. The professional might worry that the evaluation would affect his opportunities for advancement. Such anxiety might be stressful for the participants. As the logging tool we used recorded all users' actions and errors on quite a detailed level, this stress might be worse than in experiments conducted without such a logging tool.
- *Social risk*: The researcher was not involved in the teaching of the participants of the first study, but the subject of the second study was a former MSc student of the researcher. The participants in both studies might feel pressure to be positive about the technology under evaluation because they were being paid for the participation.

## Minimizing the Risks

To minimize the risks described above we informed the participants about the benefits and risks of participating in the studies (informed consent), compensated them for any inconvenience they might have experienced, and assured them that their confidentiality would be maintained.

All the participants were volunteers. At the beginning of both studies we explained to them the goals and procedure of the study and what kind of data was collected by the logging tool. They were provided with a short period of training (15 minutes) at the beginning of the experiment. The participants also had access to the documentation describing the experiment and tools and were permitted to ask for help during the studies. The participants were paid for their participation and motivated by the possibility of learning a new tool and gaining experience from participating in a usability study. We assumed that this might be viewed as adequate compensation for the inconvenience. At the beginning of the studies we assured the participants that the data collected would be treated confidentially. During the studies we tried to emphasize that we were evaluating the tools, not the participants.

But, did we do enough? Did we really succeed in minimizing risks for the participants? The participants in the first study were asked to express their agreement with statements addressing the experiment by selecting a number on a seven-point scale. Table 1 presents the results. Most of the participants found it interesting and useful to participate in the studies and were satisfied with the organization of the experi-

| Variable                                                       | Mean | Med. | StDev | Min | Max |
|----------------------------------------------------------------|------|------|-------|-----|-----|
| It was useful /interesting for me to participate               | 2.1  | 2    | 1.2   | 1   | 5   |
| Organisation of the experiment was good                        | 2.6  | 3    | 1.2   | 1   | 5   |
| I was well motivated to participate                            | 2.1  | 2    | 1.3   | 1   | 6   |
| I worked differently because I knew that everything was logged | 6.1  | 7    | 1.7   | 2   | 7   |

**Table 1.** Evaluation of the experiment (score 1 – fully agree; score 7 – fully disagreed)

ment (median 2 and 3). The students were well motivated to participate in the experiment (median 2). Two students reported problems with the programming environment and one student complained about the missing help function. The students claimed that their work was not influenced by the use of the logging tool (median 7). However, some of them wrote comments that were designed to excuse their possible poor performance, such as ‘I should know Java better’ or ‘I should be more familiar with *emacs*’. They also said that they felt time pressure and that they felt as though they were doing an examination. This indicates that they wanted their results to be as good as possible and that we might need better ways of assuring them that their confidentiality would be maintained. The participant in the second study was also paid, but he said that his main motivation was ‘interest in new tools and empirical studies’. He also found it interesting to participate in the study.

## Conclusions and Future Work

Logging users’ actions is a valuable way of collecting data during usability studies. The logging tool that we used recorded all users’ actions and errors on quite a detailed level and interrupted the participants in their work. Inconvenience and psychological risk were higher than in experiments conducted without such a logging tool. We used standard means of minimizing the risks for the participants: we obtained their informed consent, paid them for participation, and assured them that their confidentiality would be maintained.

However, minimizing risk for the participants might influence the validity of the studies. The participants might have been too positive in their participation, because they were paid volunteers. In our opinion, conflicts between validity and ethics are difficult to avoid. However, we believe that it is of greater importance to treat the participants in a study ethically than it is to construct an ideal experimental environment. The subjects cannot participate in such studies without some form of compensation.

We intend to extend the functions of the logging tool, adding options for the participants to turn off data collection or to delete some data if they wish to do so. This would allow the participants to control the logged data and thus provide them with autonomy. (Sieber, 2001).

## References

- IEEE-CS, & ACM. (1999). *Software Engineering Code of Ethics and Professional Practice*. Retrieved March, 2003 from <http://onlineethics.org/codes/softeng.html>
- Karahasanovic, A., & Sjøberg, D. I. K. (2001, 5–7 September). *Visualising Impacts of Database Schema Changes – A Controlled Experiment*. In the 2001 IEEE Symposium on Visual/Multimedia Approaches to Programming and Software Engineering, 358–365, Stresa, Italy.
- Karahasanovic, A., & Sjøberg, D. I. K. (2002, October 7–8). *Visualising Impacts of Change in Evolving Object-Oriented Systems: An Explorative Study*. In the International Workshop on Graph-Based Tools Gra BaTs’02, 22–31, Barcelona, Spain.
- Karahasanovic, A., Sjøberg, D. I. K., & Jørgensen, M. (2001, May 20–23). *Data Collection in Software Engineering Experiments*. In the Information Resources Management Association International Conference, Software Engineering Track, 1027–1028, Toronto, Ontario, Canada.

## Is it Ethical to Log Users' Actions

Onlineethics.org. (2001). *The Online Ethics Center Glossary*. Retrieved March, 2003 from <http://onlineethics.org/glossary.html>

Sieber, J. E. (1992). *Planning Ethically Responsible Research*. Newbury Park: Sage.

Sieber, J. E. (2001). Protecting Research Subjects, Employees and Researchers: Implications for Software Engineering. *Empirical Software Engineering*, 6, 329–341.

Singer, J., & Vinson, N. (2001). Why and How Research Ethics Matters to You. Yes, You! *Empirical Software Engineering*, 6, 287–290.

## Biography

**Amela Karahasanovic** is a postdoctoral fellow working in the Software Engineering Group of the Simula Research Laboratory. Her research interests are in the area of software engineering technologies and their empirical evaluation.