There are a large number of research questions on collaborative learning that cannot easily be answered by simply collecting and comparing quantitative data on the performance of individual and collaborative learners. More often than not, critical information is available only if access to a detailed recording of the students’ collaborative work is available. For certain research endeavors in the collaborative learning domain, we need to know how the students organized their task, what roles the students played, and how they participated. Videotaping is an example of a technology that has been used to investigate collaborations within the workplace. But there are some difficulties with using this approach for the studying of collaborative learning, not the least of which is the extensive time-cost of collecting and transcribing video data. We argue that groupware applications are an ideal platform for experimental investigations of collaborative learning.

At Brandeis we have been developing principles, tools, and methods for cognitively engineering groupware systems that support online collaboration. One part of this project is to develop a toolkit that enables the rapid development of groupware applications that can be used as experimental platforms. A key component of the groupware systems that are generated is that a complete transcript of all the online activity is automatically captured in a form that is replayable by an analyst using a replay device (that is created as the system is developed). Students at Brandeis have successfully used this toolkit in a HCI class to produce a working groupware system; each team of students had 28 days to write the code. Another part of our research project is to develop discourse analysis techniques for modeling the online collaborative work of users and the cognitive load it entails for each of the participants. These techniques have been taught in a class on Computational Cognitive Science.

In this poster, we will present some of the details of an experimental study of collaborative learning that we are currently conducting. Some of the questions about collaborative learning that we want to investigate are:

- How does the amount and type of participation affect individual learning?
- What do the participants talk about (i.e. which aspects of the activity do they spend the most effort on)?
- How do the participants organize their collaboration?
- How closely do the participants work together?

Corresponding to each one of these questions are significant hypotheses about the role of participation and/or explanation in collaborative learning tasks.

Our study compares the performance of individuals and pairs of students (with little or no prior programming experience) as they learn to draw figures using JScheme. As a part of this study, we constructed a platform (GrewpTool) for collaborative programming that has been used to support several kinds of classroom related activities. In our experimental study, GrewpTool collects, in a replayable transcript, the representational work of individual subjects and all of the communication between paired subjects. The participants in our experiment complete both a pre and a post test, whose score difference indicates how much they learn.

Our study has produced an enormous amount of data for analysis. The replay of an individual session of collaboration is one of the tools available for analyzing the interactional data. Given this tool, extracting specific and accurate answers to questions about participation is feasible, but the data is not easily codable and the task is labor intensive.

We have developed some automatic methods for analyzing the interaction that can be used to guide the ethnographic analysis of the subjects’ online behavior. Each of these representations is relevant to answering questions of the sort listed above. We will show automatically generated representations that depict how close each pair worked together, how they organized their collaboration, the type and amount of each subject’s participation, and the content of their conversation. Each of these representations can be created because all of the subjects’ participation is mediated by the computer and therefore automatically recorded and transcribed. Given this analysis of the data, it is possible to more selectively engage in the labor-intensive task of analyzing the replay of each session. Given these kinds of representations of data it is easier to explore the effects of explanation, participation, and organization on learning.

Acknowledgments

This work was supported by the National Science Foundation under Grant EIA-0082393, and by the Office of Naval Research under Grant N000-14-02-1-0131.