Comparing "in the Wild" Studies With Laboratory Experiments: A Case of Educational Interactive Tabletops

Bertrand Schneider, Stanford University, schneibe@stanford.edu
Consuelo Valdes, Wellesley College, cvaldes@wellesley.edu
Kelsey Temple, Wellesley College, ktempel@firstclass.wellesley.edu
Chia Shen, Harvard University, chia_shen@harvard.edu
Orit Shaer, Wellesley College, oshaer@wellesley.edu

Abstract: In this paper, we describe our attempt at implementing an interactive learning activity on a multi-touch tabletop in an actual college classroom. In our previous in-lab experimental study, we found that this tabletop learning activity outperformed the same learning task on paper. Students had a higher quality of collaboration and learnt more when interacting with an interactive multi-touch tabletop display. Here we compare those results with an "in the wild" implementation of our system. We found that learning gains were significantly lower in the classroom compared to a laboratory setting. Our paper investigates this difference and suggests reasons for such a discrepancy.

Introduction

The introduction and increasing availability of multi-touch, high-resolution displays in the form of handhelds, tabletops, and whiteboards, open the opportunity to consider these technologies as a prominent alternative to current learning technologies. Given the potential of these novel technologies, numerous research prototypes have been developed by Human-Computer Interaction researchers, exploring how these emerging technologies will impact education. However, to date, only a small fraction of prototypes have materialized outside of research lab settings and studied in actual classroom settings. A foundational work in this domain was presented by Brown (1992), who introduced *Design Experiments* as complex interventions implemented in educational settings. This approach generated controversy due to the large number of confounding variables present in classrooms. However, Brown was able to identify the fundamental tensions between a natural and artificial setting. Many researchers followed her example and recognized the importance of testing educational interventions in an ecological way. However, to date, little research has been devoted to investigating the strengths and limitations of utilizing interactive surfaces in college-level learning.

Comparing Laboratory and Classroom Experiments

Phylo-Genie, a learning activity to help college-level students grasp basic concepts in phylogenetics, is implemented on an interactive multitouch tabletop display (Microsoft Surface™). This work has been previously presented at the ACM SIGCHI conference (Schneider et al., 2012). In particular, Phylo-Genie provides students opportunities to build their own phylogenetic trees through a learning scenario.

In this section we summarize the two studies that we conducted to evaluate Phylo-Genie. The first study, a controlled experiment, was completed in a laboratory setting with college-level students; the second, in-class study, was conducted in a college classroom.

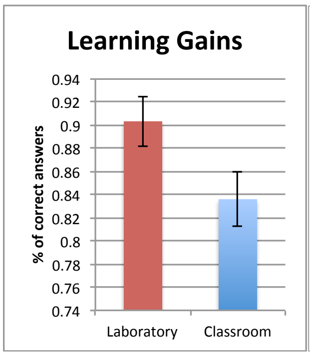
<u>Participants</u>: *controlled experiment:* 28 undergraduate and graduate students (28 female, 28 male, average age = 21.28, SD = 3.70) volunteered to participate in the study, composing 14 dyads. *Classroom setting:* 32 students, enrolled in an introduction to biology course in a women college (all females), participated in our study. None of the participants had received college-level instruction in evolutionary biology before the study.

<u>Materials</u>: The materials used in both studies were similar. A post-test was provided by a university professor who teaches introductory phylo-genetics to college-level students. We rated the quality of collaboration among dyads by using Meier's rating scheme (Meier, Spada & Rummel, 2007).

<u>Procedure</u>: We conducted the experiment in two private laboratory spaces for the controlled experiment, and in the classroom for the second study. Subjects received a 5-minutes introduction to phylogenetics before the Phylo-Genie activity. They then went through all the steps of the learning scenario and completed a post-test. Each session was approximately 60 minutes long. Due to limitations in hardware availability, some students had to work in groups of three in the classroom condition.

<u>Results</u>: we found that students in the laboratory setting outperformed students in the classroom on the learning test: $F(1,58) = 4.32$, $p < 0.05$ (mean=0.9, SD=0.11 for the laboratory, mean=0.83, SD=0.13 for the classroom).

We also rated the quality of collaboration of each dyad participating in the tabletop activity in the classroom and in the laboratory according to Meier's, rating scheme (Meier, Spada & Rummel, 2007). We found that overall, students had a better quality of collaboration in the classroom: F(1,16) = 4.26, p = 0.056, eta squared = 0.21 (for the classroom, mean = 21.00, SD = 8.68; for the laboratory, mean = 13.67, SD = 6.19).



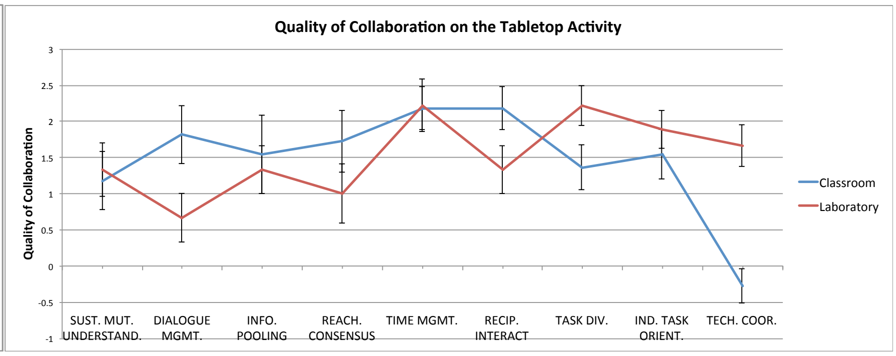


Figure 1. Quality of collaboration in each group and setting (based on Meier's rating scheme).

The sub-dimensions *dialogue management* (F(1,16) = 16.67, p = 0.001), *reaching consensus* (F(1,16) = 5.50, p = 0.032), *reciprocal interaction* (F(1,16) = 7.27, p = 0.016), *task division* (F(1,16) = 3.16, p = 0.094, eta squared = 0.035) and *technical coordination* (F(1,16) = 4.26, p < 0.001) showed a statistical difference (Figure 1). Interestingly, only *task division* was correlated with a positive learning gain across our two groups.

An interesting detail is that the five lowest scoring students in the classroom (in terms of learning gains) were all in different triads; compared to the mean learning-gain score of the classroom tabletop condition, two of these students were below two standard deviations, one was below one standard deviation and the remaining two were within one standard deviation of the mean. If we remove the two outliers (lowest scoring students) from the analysis, the difference in terms of learning gains between the classroom and the lab conditions become non-significant: F(1,56) = 2.45, p = 0.123. We also have more data concerning students' territoriality and discourse, but due to space constraint we cannot describe them here.

Discussion

Our initial results suggest that triads seem to be more likely to have one "weak" student in the group. We believe that these students were not less able, but they may have been excluded (to some extent) from the rest of the group. The "task division" sub-dimension of the collaborative rating scheme was able to detect this difference between dyads and triads. More fine-grained studies should look at how students in triads become disengaged or get excluded from the learning activity.

Our preliminary results suggest that *hardware availability* may have caused students in the classroom to learn less than in a laboratory setting. By having students work in triads instead of dyads, groups were more likely to have a "free rider" (e.g. a student who does not engage in the activity). This effect is described by Salomon & Globerson (1989), where implicit coalitions are detrimental to collaborative learning. From a technical standpoint, this is a critical issue. Interactive tabletops are expensive, and most schools cannot afford to buy multiple devices. Consequently, researchers should pay a special attention to the ratio of interactive tabletops per students when implementing this kind of technology in the classroom. For future work, we plan to conduct more in-depth analyses of students' interaction to unpack (and suggest ways to prevent) group coalition in the classroom.

References

Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences*, *2*(2), 141–178.

MacKinnon, K. (2012) Context matters: The value of analyzing human factors within educational contexts as a way of informing technology-related decisions within design research. *ijcscl* **7** (3)

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *ijCSCL, 2*(1), 63–86.

Shaer, O. et al., The Design, Development, and Deployment of a Tabletop Interface for Collaborative Exploration of Genomic Data, *International Journal of Human-Computer Studies,* 2012.

Schneider, B., Strait, M., Muller, L., Elfenbein, S., Shaer, O., Shen, C. (2012). Phylo-Genie: Engaging Students in Collaborative 'Tree- Thinking' through Tabletop Techniques. In Proc. *ACM Conference on Human Factors in Computing Systems (CHI'12),*ACM Press.