<u>“Students’ Motivation and Learning of Science in a Multi-User Virtual Environment”</u>

Chris Dede, Jody Clarke, Diane Jass Ketelhut, Brian Nelson, and Cassie Bowman
Harvard Graduate School of Education

Abstract: This NSF-funded project utilizes graphical multi-user virtual environments (MUVEs) as a vehicle to study (1) classroom-based situated learning and (2) the ways in which virtual environments may aid the transfer of learning from classroom contexts into real world settings. In the project’s River City curriculum, teams of middle school students are asked to collaboratively solve a simulated 19th century city’s problems with illness, through interaction with each others’ “avatars,” digital artifacts, tacit visual and auditory clues, and computer-based “agents” acting as mentors and colleagues in a virtual community of practice. In this paper, we provide an overview of results from a large-scale implementation of the River City environment and curriculum in Spring 2004. Our findings show that students and teachers were highly engaged, that student attendance improved, that disruptive behavior dropped, and that interesting patterns are emerging about which students do best under our various pedagogical conditions.

Keywords: technology, situated learning, inquiry, virtual environments, gender.

OBJECTIVES

A major goal for education in the 21st century is to create scientifically literate citizens who are able to think critically, make sense of complex data, and solve problems (AAAS, 1993; NRC, 1996). Research suggests that, if all students are to become scientifically literate, science instruction must convey greater engagement and meaning to them. To achieve this, science instruction in secondary schools must provide students with opportunities to explore the world, to apply scientific principles, to sample and analyze data and to make connections among these explorations, their personal lives, and their communities. However, given the constraints of classroom settings, real world data collection is challenging to orchestrate. Due to safety issues and lack of equipment, laboratory experiments are also difficult to conduct. It is no surprise that educators report teaching higher order inquiry skills (such as hypothesis formation and experimental design) is among the most difficult challenges they face with students who have a history of low achievement in and engagement with science.

With NSF funding, we are creating and studying graphical multi-user virtual environments (MUVEs) to enhance middle school students' motivation and learning about science and society (http://muve.gse.harvard.edu/muvees2003/). MUVEs enable multiple simultaneous participants to access virtual contexts, to interact with digital artifacts, to represent themselves through “avatars,” to communicate with other participants and with computer-based agents, and to enact collaborative learning activities of various types. Our “River City” MUVE is centered on higher order scientific inquiry skills, as well as on content related to national standards in biology and ecology.

In this paper, we provide an overview of results from May, 2004 of a large-scale implementation of the River City environment and curriculum. We conducted implementations of more than 1000 students in Boston- and Milwaukee-area classrooms, with high proportions of ESL and free-and-reduced-lunch pupils.

THEORETICAL FRAMEWORK

Brown, Collins, & Duguid (1989) and Lave & Wenger (1991) define situated learning theory as embedded within and inseparable from participating in a system of activity deeply determined by a particular physical and cultural setting (Chaiklin & Lave, 1993; Lave, 1988). The unit of analysis is neither the individual nor the setting, but instead the relationship between the two, as indicated by the student's level of participation in the setting (Barab & Plucker, 2002). Studies of apprenticeship in "communities of practice" (moving from newcomer to expert within a sociocultural structure of practices) are a central construct for situated learning (McLellan, 1996; Kirshner & Whitson, 1997; Wenger, 1998; Wenger, McDermott, & Snyder, 2002).

In essence, situated learning requires authentic contexts, activities, and assessments coupled with guidance based on expert modeling, situated mentoring, and legitimate peripheral participation. Greeno (1997) indicates that the power of situated learning is derived from a person learning to solve problems as part of a community in the authentic context confronting these challenges, a difficult environment to develop in a classroom. Part of the promise of MUVEs is their capability to create immersive, extended experiences in the classroom with problems and contexts similar to the real world.

The River City MUVE curriculum is based on students participating in an elaborate context modeled on the real world, interacting with novices and experts who are part of its culture. Learners actively investigate multivariate problems with aid from community members and mentors with various types of expertise. To study situated learning in this implementation, we utilize a virtual expert to model inquiry skills and then coach students while they learn these techniques (Griffin, 1995).

NATURE OF THE INTERVENTION

The River City curriculum focuses on skills of hypothesis formation and experimental design, conveyed via standards-based content in biology and ecology. Students gain knowledge through immersive simulations, interactive virtual museum exhibits, and "participatory" historical situations. Students learn to behave as scientists while they collaboratively identify problems through observation and inference, form and test hypotheses, and deduce evidence-based conclusions about underlying causes.

The River City virtual "world" consists of a city with a river running through it; different forms of terrain that influence water runoff; and various neighborhoods, industries, and institutions, such as a hospital and a university. The students themselves populate the city, along with computer-based agents, digital objects that can include audio or video clips, and the avatars of instructors (Figure 1). Content in the right-hand interface-window shifts based on what the participant encounters or activates in the virtual environment (Figure 2).

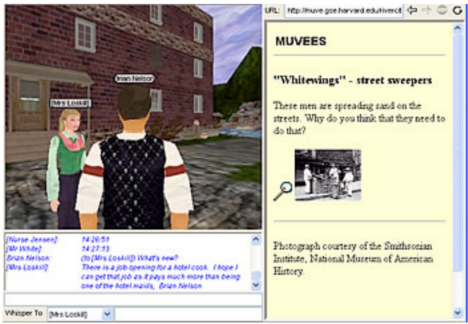


Figure 1

Figure 2

In River City, students work in teams to develop hypotheses regarding one of three strands of illness in the town (water-borne, air-borne, and insect-borne). These three disease strands are integrated with historical, social and geographical content, allowing students to experience the inquiry skills involved in disentangling multi-causal problems embedded within a complex environment. At the end of the project, students compare their research with other teams of students in their class to delineate the many potential hypotheses and causal relationships embedded in the virtual environment.

For this implementation, we developed three variations of the River City curriculum in order to start exploring the type of learning best supported by MUVEs. Variant GSC centers on a guided social constructivist (GSC) model of learning-by-doing, in which guided inquiry experiences in the MUVE alternate with in-class interpretive sessions led by the teacher. Variant EMC shifts the learning model to a situated pedagogy with expert modeling and coaching (EMC), based on expert agents embedded in the MUVE. Our third "control" condition utilizes a curriculum in which the same content and skills were taught in equivalent time to comparable students in a paper-based format without technology, via a guided social constructivist-based pedagogy. This type of control curriculum enables our research to focus on the strengths and limits of MUVEs as well as the types of pedagogy best supported by this medium.

The research questions that guided this study are:

1. When compared to the "control" version, what types of significant gains in motivation and learning does version GSC produce?

2. What differences in student outcomes between versions GSC and EMC seem attributable to their contrasting pedagogies?

RESEARCH DESIGN

Methods and Procedure

We conducted large-scale implementations with 11 teachers and more than 1000 students in Boston- and Milwaukee-area classrooms, with high proportions of ESL and free-and-reduced-lunch students. Data for approximately 300 students was analyzed as a representative sample. To control for threats to validity, the two computer-based variants (GSC and EMC) were randomly assigned to students within each classroom, with teachers instructed to minimize cross-contamination of treatments. A third, paper-based control treatment was randomly assigned to whole classes. Each teacher offered both the computer-based treatments and the control.

Both quantitative and qualitative data were collected from students and teachers over the three-week implementation period. The students were administered an affective measure that was adapted from three different surveys, Self-Efficacy in Technology and Science (Ketelhut, 2004), Patterns for Adaptive Learning Survey (Midgley, C. 2000), and the Test of Science Related Attitudes (Fraser, 1981) pre- and post-intervention. This modified version has scales to evaluate students' efficacy of technology use (videogame, computer, chat), science efficacy, thoughtfulness of inquiry, science enjoyment, career interest in science, etc. To assess understanding and content knowledge (science inquiry skills, science process skills, biology), we administered a content test, (with sections modified from Dillashaw and Okey, 1980), pre- and post-intervention. In addition, log files of individual student activity in the MUVE were captured for all students in test classrooms. After designing and conducting their experiments, students in both the control and River City treatments were asked to write letters to the Mayor of River City in which they discussed their hypothesis, experimental design, results, and recommendations for solving the city's health problem.

We conducted semi-structured interviews with 12 students (six boys and six girls) pre-, during, and post-intervention. The students were chosen by their teacher and represented both low and high achievement. Interviews were conducted in the school during the students' free period. All interviews were audio recorded and transcribed verbatim.

The teachers participated in an online professional development program, focused on content review, alternative pedagogical strategies based on different theories of learning, facilitation strategies while students are using the MUVE, and interpretive strategies for leading class discussions. They collected demographic data and rated their expectations of students' successes and motivation with the project, and responded to a pre and post questionnaire regarding their methods, comfort with technology, and reflections on using the MUVE in their science class.

Data Analysis

The quantitative data were analyzed with SAS. We ran descriptive statistics and correlations and performed multilevel modeling (MLM) techniques. A significance level of $p \leq .05$ was used; and checks for linearity, normality and homoscedasticity were performed at various intervals. No clear violations were noted.

Open coding techniques (Strauss and Corbin, 1998) were used to code all of the interview transcripts. Cross-case analyses were conducted to compare patterns and themes that emerged. We then compiled cases of individual students to illustrate these themes.

RESULTS

In this section we briefly describe some of our key findings.

Biology Content Results

Students in the River City experimental treatments improved their biological knowledge by 32%-35%. Control students also improved, but by only 17%.

Inquiry Content Results

When using survey questions to assess inquiry, improvements were seen across the board for knowledge and application of scientific processes; control students improved slightly more than the other two groups: 20% for the control, 18% for the GSC group and 16% for the EMC group.

Letters to the Mayor

We believe it is difficult to measure inquiry with a multiple-choice test. Therefore, we also analyzed the letters to the mayor that students wrote as a final understanding performance for evidence of inquiry learning. Since detailed comparison of the letters from the River City curriculum and the control curriculum may not be productive, as it is harder to conduct an experiment on paper, we looked for similar demonstrations of student understanding of the processes of inquiry and for motivation. The letters written for the control curriculum often: were much shorter in length, did not demonstrate motivation or engagement, did not mention the experiment, and did not explicitly recognize the interconnectedness of the chosen problem with other possible causes of the larger problem. Analysis of the letters' evidence of inquiry found that students taking part in the MUVE-based curriculum earned scores more than double that of their paper-based control peers, on average ($p<.01$).

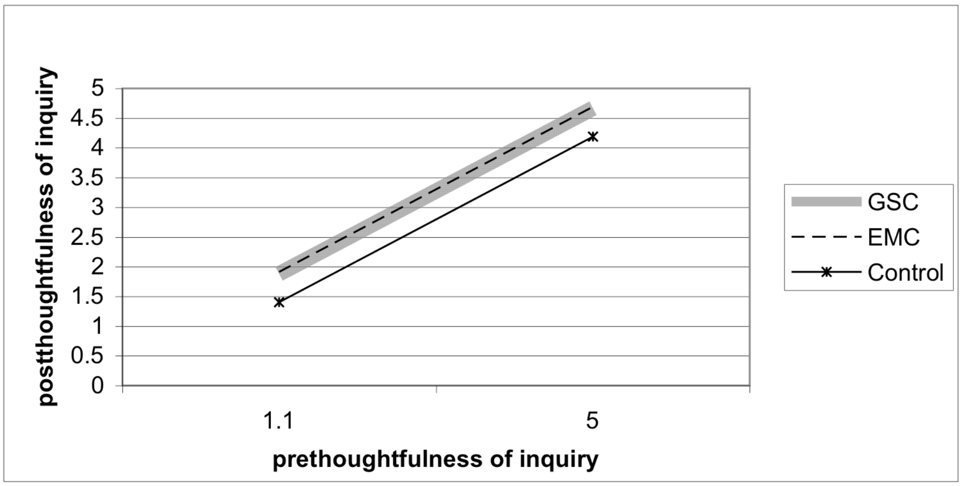
Further analysis of students' letters to the mayor of River City suggest that students demonstrate an understanding of the process of inquiry that was not well captured in the science inquiry post-test measures. For example, students who scored low on the science inquiry post-test wrote letters that were of similar quality to those written by students who scored higher on the post-test. As one illustration, in their letters low-performing content students matched the high-performing content students around criteria of stating an opinion regarding the cause of the problem and/or the outcome of the experiment. In addition, in their letters both low- and high-performing students demonstrated a clear causal relationship between the problem and the reason(s) for the problem.

Interestingly, more of the lower-performing test students met the criteria of providing suggested interventions or further research than students who scored higher on the inquiry test questions. This suggests that the complexity of the MUVE treatment creates intricate patterns of learning more appropriately measured with an authentic activity, such as writing an experimental report.

Affective Results

We were also interested in characteristics that promote scientific interest and inquiry. For example, on our affective measure test, we measured *thoughtfulness of inquiry,* a measure of metacognitive awareness. Student scores on this subscale on the post-survey were significantly higher ($p<.01$) on average for River City students, in comparison to the scores for students in the control group (see figure 4 below).

Figure 4: Effect of treatment on post thoughtfulness of inquiry controlling for pre-thoughtfulness of inquiry (n=330)



For example, River City students scoring an average of 1 (strongly disagree) on the scale of 1-5 for the pretest were associated with scores of 1.8-1.9 on the posttest, nearly double their starting average score. Students in the control group also improved, on average, but only to 1.3. Another subscale measured *interest in a scientific career*; the gain in interest in science careers was 5% higher for students who had taken part in the River City curriculum than for those who had completed the control curriculum—a substantial gain for a 2-week implementation.

Engagement in River City

Three pieces of evidence support the high level of engagement we saw in the classroom. First, students and teachers in interviews and focus groups repeatedly expressed interest in using River City again. Teachers also offered positive reports about student engagement in the project. For example, one teacher reported that "Students seem to learn more deeply about science and problem solving in the simulation or manipulation setting than in traditional book education, as evidenced by their class discussions" (Galas & Ketelhut, in press). A number of students reported that it was the first time they enjoyed science class (Clarke & Dede, 2005).

The second piece of evidence indicating engagement came from the absentee records. One of our school districts had absentee rates approaching 50% during the time frame of our implementation. For the single participating teacher in that district, absentee rates decreased by

35% from the first to last week of the project. We are now in the process of analyzing absentee data from other sites.

The final piece of evidence on engagement comes from a swearing monitor we installed to allay concerns by teachers that allowing students to chat online during school would result in the use of inappropriate language. This turned out to be a very minor problem that nearly disappeared by the end of the project. There were only about 70 instances of inappropriate language in the first two days in the world from thousands of interchanges by over one thousand students. By the last days in the world, this had dropped to 13 instances of inappropriate language out of thousands of interactions. These interactions also show that students are building fluency in virtual communication and expression, important skills for the 21st century workplace.

CONCLUSION & IMPLICATIONS

MUVE research is in its infancy. To date, we are the only research group studying whether MUVEs aligned with core curriculum, national standards, and 21st Century Skills increase educational effectiveness in middle school classrooms. Our project is showing that MUVE-based curricula can teach standards-based biological content infused with complex inquiry skills better than good traditional approaches do. Our findings show that students learned biology content, that students and teachers were highly engaged, that student attendance improved, that disruptive behavior dropped, that students were building 21st century skills in virtual communication and expression, and importantly, that using this type of technology in the classroom can facilitate good inquiry learning.

Our work is helping the field understand the strengths and limits of this medium for learning, an important topic at a time when many students spend much time in virtual environments outside of school and when strong claims are made about the educational effectiveness of online games.

REFERENCES

American Association for the Advancement of Science. (1993). Benchmarks for science literacy. New York, NY: Oxford University Press.

Barab, S. A., & Plucker, J. A. (2002). Smart People or Smart Contexts? Cognition, Ability, and Talent Development in an Age of Situated Approaches to knowing and Learning. *Educational Psychologist, 37*(3), 165-182.

Brown, J. S., Collins, A., & Duguid. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher, 18*(1), 32-42.

Chaiklin, S., & Lave, J. (1993). *Understanding practice: Perspectives on activity and context.* New York: Cambridge University Press.

Clarke, J. & Dede, C. (2005). Making Learning Meaningful. Paper presented at AERA, Montreal.

Dillashaw, F. G., & Okey, J. R. (1980). Test of integrated process skills for secondary science students. *Science Education, 64*(5), 601-608.

Fraser, B. (1981). *TOSRA: Test of Science Related Attitudes*. Australian Council for Educational Research, Hawthorne, VIC.

Galas, C & Ketelhut, D. (in press). River City, The MUVE. *Leading and Learning with Technology*.

Greeno, J. G. (1997). On Claims That Answer the Wrong Questions. *Educational Researcher, 26*(1), 5-17.

Griffin, M. M. (1995). You Can't Get There from Here: Situated Learning, Transfer, and Map Skills. *Contemporary Educational Psychology, 20*, 65-87.

Ketelhut, D. (2004). Self-Efficacy in Technology Survey. Qualifying Paper presented to Harvard Graduate School of Education. Cambridge, MA.

Lave, J. (1988). *Cognition in practice***.** Cambridge, UK: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

McLellan, H. (Ed.). (1996). Situated learning perspectives. Englewood Cliffs, NJ: Educational Technology Publications.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., Gheen, M., Kaplan, A., Kumar, R., Middleton, M. J., Nelson, J., Roeser, R., & Urdan, T. (2000*). Manual for the Patterns of Adaptive Learning Scales (PALS)*, Ann Arbor, MI: University of Michigan

National Research Council. (1996). National science education standards. Washington, DC.

Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Thousand Oaks, CA: Sage.

Wenger, E. (1998). Communities of practice. New York: Cambridge University Press.

Wenger, E., McDermott, R., & Snyder, W. (2002). Cultivating Communities of Practice. Cambridge, MA: Harvard Business School Press.