Designing for Motivation and Usability in a Museum-based Multi-User Virtual Environment

Chris Dede & Diane Ketelhut
Graduate School of Education
Harvard University
Cambridge, MA 02138
Chris_Dede@harvard.edu
ketelhdi@gse.harvard.edu

The Multi-User Virtual Environment Experiential Simulator (MUVEES) is the result of a two-year National Science Foundation (NSF) funded research project. The intent was to create and evaluate graphical multi-user virtual environments (MUVEs) that use digitized museum resources to enhance middle school students' motivation and learning about science and its impacts on society. MUVEES was designed to take advantage of known affordances of MUVEs where multiple participants engage in collaborative learning activities by accessing virtual environments synchronously and by interacting with digital artifacts. Participants represent themselves through *avatars*[1] and communicate with other participants and computer-based agents. Initially, MUVEs were based only on textual descriptions; now, many MUVEs are graphical in nature or use graphics to enhance textual descriptions. Our project's educational environments are extending current MUVE capabilities in order to study the science learning potential of immersive simulations, interactive virtual museum exhibits, and "participatory" historical situations (http://www.gse.harvard.edu/~dedech/muvees/). To accomplish this, we have built our own MUVE shell based on the Sense8 WorldToolKit (http://www.sense8.com/).

Our MUVEs are not stand-alone learning environments, but instead are designed for use in a classroom context, supplemented by conventional instructional activities such as textbooks and teacher-led discussions. Hard-copy outputs, including both lab-notebook answers and written letters to the “mayor,” provide methods for assessment. Teachers use these student products and "cognitive audit trails" of activities in the MUVE to determine when to promote a learner to the next level in a unit. Our virtual learning environments function on bandwidth as low as a 56K modem and on Windows desktop computers with at least 450 Megahertz processing speed and 64 Megabytes of RAM, well within the specifications of machines currently being purchased by schools.

Our project is studying how the design characteristics and usability of MUVE learning experiences affect middle school students' motivation and educational outcomes. We are focusing on the extent to which guided inquiry and mentoring using digitized museum artifacts in virtual contexts can aid pupils' performance both on classroom assessments and on high-stakes

[1] Graphical representation of the player in the virtual world, controlled by the player

tests related to national science standards. In this research, we are studying pedagogical strategies that foster strong learning outcomes across a wide range of individual student characteristics (e.g., prior experience with computers, knowledge about and interest in science, gender, ethnicity, linguistic proficiency in English). Interaction in the MUVEs involves substantial reading and writing, so we are providing all student materials in both English and Spanish to study the effects of this intensive textual interaction in enhancing multi-lingual communication and literacy.

In particular, we are working to dramatically improve the educational outcomes of the bottom-third of students, pupils who even by middle school often have given up on themselves as learners. These students are disengaged from schooling and typically are difficult to motivate even by good teachers using inquiry-based pedagogy. We are studying whether educational MUVEs with deep content and challenging activities that resemble the entertainment and communication media these students use outside of school can reengage them in learning. Our research is also examining the implementation process needed to successfully integrate MUVEs in typical classroom settings. Our long-range goal is to develop "replacement" units that—through simultaneously developing students' knowledge in science, mathematics, history, and social studies—enable the teaching staff to utilize inquiry methods capable of covering the same span of content in a similar amount of time required by presentational teaching of this material one topic at a time. Given current educational "reforms" based on an overcrowded curriculum driven by high-stakes, low-level tests, this is essential for any scalable curricular intervention.

At various times, partners in this research were Harvard's Graduate School of Education, the Virtual Environments Lab at George Mason University, the Smithsonian's National Museum of American History (NMAH), and Thoughtful Technologies, Inc. The development team included education researchers, science educators, instructional designers, computer scientists, museum archivists and exhibit designers, graphic artists, scientists, and middle-school science teachers from both public and private schools. We are documenting our design process to generate heuristics generalizable to other groups developing educational multi-user virtual environments. This study reports on a “River City” curriculum unit on which we conducted pilot implementations in public school classrooms in Boston, MA.

RESEARCH QUESTIONS

Our research questions revolved around using MUVEES to understand the affordances of MUVEs in the classroom:

- Does using a MUVE improve learning for all students, particularly those unmotivated or alienated by more traditional methods?

- What are the implementation issues in using a MUVE?
- Can technology such as a MUVE improve motivation and interest in science, especially for students in the bottom third of academic achievement?

This pilot implementation examined these research questions. Its promising results demonstrate the need for additional studies into how MUVEs can not only further student learning, but also can facilitate research into pedagogical practices and theories.

DESIGN AND DEVELOPMENT

Our initial design started with problem areas identified by science teachers. They were interested in experimenting with an intervention that addressed difficult parts of the curriculum, not relatively easy material for which well functioning instructional strategies were already available. For middle school, the teachers identified experimental design as the most difficult concept for students to learn and the one for which teachers were most eager to find an alternative instructional method. In particular, an activity in which students behaved as scientists while they identified a problem through observation and inference and then formed and tested hypotheses was seen as potentially very valuable. To minimize challenges in professional development and implementation, we developed hard-copy “lab notebooks” that students use throughout the unit to guide their inquiry; these provide a familiar way for teachers to assess student progress through the unit and learning of the science content.

We have developed two prototype middle school science curriculum units, each based around the national science standards, content typically covered in the middle school science curriculum, and the types of investigative process skills necessary for students to do an independent science fair project. Each unit consists of a multi-user virtual environment with virtual contexts and digital artifacts that directly and tacitly guide learner investigations. Student materials for one of the units are in both English and Spanish. Our first curriculum prototype (the River City Unit) centers on content in biology and ecology and is the topic of this paper. The second curriculum prototype (the Bicycle unit) focuses on the physical and material sciences and is centered on problems of bicycle design related to Newtonian motion.

The River City unit is based on students collaboratively investigating a virtual “world” consisting of a city with a river running through it, different forms of terrain that influence water runoff, houses, industries, and institutions such as a hospital and a university. The learners themselves populate the city, along with non-player characters (NPCs), digital objects that can include audio or video clips, and computer-based agents. River City contains over fifty digital objects from the Smithsonian’s collection, plus "data collection stations" that provide detailed information about water samples at various spots in the world. Figure 1 is a screen shot depicting a "birdseye" view of this city. (All figures are at the end of the paper.)

River City is typical of the United States in the late nineteenth century; Figure 2 depicts how we use museum artifacts to illustrate building exteriors and street scenes from that period in history. Content in the right-hand interface-window changes based on what the participant encounters or activates in the virtual environment. Dialogue is shown in the text box below these two windows; members of each team can communicate regardless of distance, but in-team dialogue is displayed only to members of that team. Participants can choose to interact either within their team or with all participants in the MUVE at that time. To aid their interactions, participants also have access to one-click interface features that enable the avatar to express (through stylized postures and gestures) emotions such as happiness, sadness, and anger. These interface features also allow looking upward or downward and seeing the world from a first-person perspective or from behind one's own body in a third-person viewpoint.

Multiple teams of students can access the MUVE simultaneously, each individual manipulating an avatar through their computer. In our implementations, the class is divided into teams of two to four students, which are "sent back in time" to this virtual environment. The lab notebook that the student teams use asks the class to help the city solve its environmental and health problems, which are directly related to middle school science content. To accomplish this, the students must collaborate to share the data each team collects. Beyond textual conversation, students can project to each other "snapshots" of their current individual point of view (when someone has discovered an item of general interest) and also can "teleport" to join anyone on their team for joint investigation. In-group collaboration is visible only to other group members. Each time a team reenters the world, several months of time have passed in River City, so that learners can track the dynamic evolution of local problems.

Learners are engaged in a “participatory historical situation” in which they can apply tools and knowledge from both the past and the present to resolve an authentic problem. In this “back to the future” situation, students’ mastery of 21st century classroom content and skills empowers them in the 19th century virtual world. Figure 3 shows the 19th century laboratory equipment representative of what students can use to aid themselves in River City. An eventual goal of our work, not currently implemented, is the capability of interaction with some digital artifacts. For example, were a learner to double-click on a piece of apparatus, such as a microscope, the corresponding avatar would look through a virtual microscope; and the image from the microscope slide would appear in the right-hand interface-window.

Figure 4 presents water quality data from one of eleven water-sampling stations in River City. Potentially, at the high school level, this could link to virtual laboratory equipment applications such as the irYdium project at Carnegie Mellon University (http://ir.chem.cmu.edu/irproject/***)***, so that students could conduct their own analyses rather than receiving the experimental results. Through data gathering, students observe the patterns that emerge and wrestle with questions such as “Why are many more poor people getting sick than rich people?” Multiple causal factors are involved, including polluted water runoff to low-lying areas, insect vectors in swampy areas, overcrowding, and the cost of access to medical care. Eventually, we plan to develop the capability for participants' avatars to become ill if they behave in ways that trigger these factors in the environment, thus increasing students' sense of immersion in the world.

Throughout the world, students encounter residents of River City and "overhear" their conversations with one another. These nonplaying characters (NPCs) are computer-based "agents" who disclose information and indirect clues about what is going on in River City [Figure 5]. As with the content displayed in the right frame, the phrases "spoken" by NPCs vary according to a student's level.

The main goal of the MUVE is to teach students the skills necessary for scientific inquiry, such as would be important in conducting investigations for a science fair project. This goal follows the National Science Education Standard A for scientific inquiry in 5-8 grades (http://www.nap.edu/readingroom/books/nses/html/6d.html). The emphasis in the MUVE is on identifying a problem, and therefore, River City has multiple lines of potential exploration. As mentioned above, there are 3 main strands of illness in River City. Some inhabitants have intestinal upsets from water contaminated by sewage. The sewage reaches the river from deliberate outflow of the flush toilets newly introduced to the wealthy homes in the mountains. This contaminated water flows downstream to the swampy lands behind the tenement homes and is used by the poor as potable water. This potential strand of investigation aligns with Standard F, Science in Personal and Social Perspectives, and demonstrates to students the complications of human activity as well as the effect of differing living conditions on disease transmission (e.g., the wealthy live in single family homes upstream, while the poor live downstream in crowded conditions promoting disease contagion). In addition, some of the city wells, used by middle class residents and hotel guests, are also contaminated from street runoff of manure. Investigating the water contamination integrates the students' knowledge of the water cycle, Standard D, Earth Science, with the historical aspects of River City.

The second disease strand is the introduction from outside of tuberculosis and the patterns that emerge of its spread throughout River City. The students' awareness of the passage of time is particularly interesting for this disease and for the last, malaria. Malaria, endemic to many areas of the U.S. in the 1800's, follows the life cycle of its carrier, the mosquito. As a result, as the students log in each day, they are able, if this problem is of interest to them, to follow the rise and fall of new cases of malaria from summer to winter and back again. These three diseases were chosen to represent three different disease vectors: water-borne, air-borne, and insect-carried, and align with Standard C, Life Science.

These three disease strands are integrated with historical, social and geographical content to allow students to experience the realities of identifying a problem, from within a content-rich environment, to investigate. Many students learn in schools the unrealistic view that there is a single right answer in science, easily discernible. In exploring River City, however, students are each guided in teams to make a unique hypothesis regarding one of many problems, based on their own interest. At the end of the project, they compare their research with other teams of students to discover the plethora of potential hypotheses and avenues of investigation available for exploration.

During their time in the MUVE, students answer questions in a Lab Notebook, which the teachers later use for assessment purposes. The Lab Notebook starts with questions that guide exploration of the environment and develop mastery of the interface, building towards later investigations that are content specific and require completing a data table based on the water samples encountered in River City. At the end, students write a letter to the mayor of River City describing the health and environmental problems they have encountered and making suggestions for improving the life of the inhabitants.

Over time, we plan to add various capabilities that may enhance participants' motivation and engagement. In addition to features listed earlier, these include:

- An avatar can gain new powers through mastering skills and knowledge. Learners who master science content and investigative skills can attain higher "levels of performance" in which they can use doors and objects that lower level participants cannot. To keep the interactions collaborative and to help all participants learn, part of attaining a level is based on helping a less adept student master content and skills through peer mentoring, thus keeping all students on close to the same levels.
- As students succeed in resolving problems in River City, new images can appear in the world (clearly labeled as pseudo-historical objects) in which the learners are shown as co-habiting the historical environment with figures from the past. These non-player-character (NPC) avatars can provide specific information about the environment and their lives, information necessary for identifying and solving problems related to the curriculum on which the unit is based. .
- Historical figures can appear as avatars to guide students in resolving the problems in River City (parallel to encountering an actor playing Thomas Jefferson on a visit to historic Williamsburg). For example, Ellen Swallow Richards (the first woman to graduate from MIT) was a pioneer in water quality treatment and is the “patron saint” of River City. Contemporary experts such as public health officials could also at times assume avatar roles in the MUVE to illustrate present day water quality issues in the U.S. and in developing countries. Such figures could provide not only mentoring, but also role models, especially for young women and minority students fighting cultural stereotypes about learning science.

- Students can travel through time to the same location at different historical eras to develop an understanding of how use of natural resources and technologies affect both the quality of life and the environment. For example, while situated in River City during the 1960s, students could gather information from a Rachel Carson-like historical figure to assess the effect of pesticides on water quality, the environment, and crop yields. They could also have access to antibiotics for treating illnesses caused by water-borne diseases.

Videogames and chat tools that incorporate some of these features seem quite motivating to many children.

METHODS, EVIDENCE, AND RESULTS

With the assistance of school administration, two classrooms that had access to the technology hardware and software required by MUVEES were identified in the Boston Public Schools for participation in this project. They were located in different middle schools, each with its own racial/ethnic mix. One was a 7th grade classroom in School A, serving a population of primarily African-American and Hispanic families. School B had a large Asian population, and we worked with a sixth grade class there. Control classrooms with similar attributes were chosen in each school. In School B, the experimental and control classes were taught by different teachers; in School A, the same teacher taught both. The control curriculum was designed to match the intervention curriculum on as many characteristics as possible other than technology. The intervention lasted for a total of three weeks. There were 45 students in the two experimental classes, and 36 in the control, evenly split by gender.

Both qualitative and quantitative data were collected from students and teachers over the three-week implementation period. Patterns for Adaptive Learning Survey (Midgley, 2000) with subsections on science interest, thoughtfulness of inquiry, motivation, collaboration, academic efficacy, technology interest, etc and a content test, (modified from Tobin, 1999), were administered to students, pre- and post-intervention (see Appendix A). In addition, demographic data and teachers' expectations of students' successes were collected. Table 1 lists the type of data that was collected, how it was collected and whether and how it has been analyzed to date.

Table 1. Categories of: Data gathered, Collection Methods, and Analysis Methods Used

Type of data	*Collection Method*	*Analyzed*
Content Knowledge: Scientific method literacy, Disease transmission, Problem identification, historical development of disease understanding	Pre-Post Student Test	Descriptive statistics, Multiple Regression Modeling broken down by overall changes and changes related to type of question (multiple choice versus short answer) and type of content covered (scientific method, disease, metacognitive);
Attitudes towards Science, student perceptions of: academic efficacy, teacher pressing for understanding, motivation, science interest, thoughtfulness of inquiry, collaboration, tech interest, seeing real world connections, attention, organizational skills, belief that science will help them in future	Pre-Post Student Survey—modified 'PALS' (survey with internal consistency reliability and validity measures conducted)	Descriptive statistics, Multiple Regression Modeling broken down by changes related to specific affective areas
Demographic Information: gender, race, ethnicity, first language in the home, SES, previous success in science, access to technology outside classroom, standardized test scores, visits to museums	1. Student reports 2. Student files 3. Teacher reports	Descriptive statistics, Multiple Regression control factors
Lab Notebook	Student Activity	Not evaluated
Log files	Software collection	Not evaluated
Observational data	Researcher reports	Not evaluated
Teacher Factors: teaching style, expectation of student performance	Teacher reports	Not evaluated
Teacher Attitudes towards technology	Pre-Post Survey	Not evaluated
Teacher Attitudes towards the MUVE	Teacher report at end	Ideas incorporated into re-design of professional development and implementation

Observational data was collected from the test classrooms throughout the project and sporadically from the control classrooms. All teachers responded to a pre and post questionnaire regarding their methods and comfort with technology. The test classroom teachers also wrote a narrative about their perceptions of the MUVE at the end of the project.

The content test covered the spectrum from recall questions to metacognitive questions. The affective measures instrument (PALS+) was divided into 12 subsections (Table 2) and was read to students to control for reading ability. While students worked on the content test at their own pace, questions of reading were answered by the administrator.

The variables we identified to help answer our research questions are shown in Table 2.

Table 2. Areas in which information was gathered during the MUVEES implementation and the source of the information (n=81)

Demographic Variables	**Content Test**	**PALS+ Affective Measures Survey**		
		Affective Area with example question	*Number of questions*	*Scale: 1 (not true at all) to 5 (very true) unless noted*
Grade	Pretest Score	Academic Efficacy Rating: *I'm certain I can master the skills taught in class this year.*	**5**	
Gender	Posttest Score	Motivation Mastery: *I do my science work because I'm interested in it.*	**6**	
Race/Ethnicity: White, African-American, Hispanic, Asian	Score on Multiple-Choice Questions only	Science Interest*: I enjoy what we do in science class.*	**6**	Responses to questions with a negative cast were reversed for scoring to give the equivalent positive rating.
Science GPA	Score on Scientific	Thoughtfulness of Inquiry: *When I look at my data, I*	**12**	

this year	Method Literacy only	*try to think about how it answers the question.*		
SES (measured by free or reduced lunch status)		Technology Interest: *I like learning to use technology.*	**4**	Responses to questions with a negative cast were reversed for scoring to give the equivalent positive rating.
Teacher Expectation of Success		Collaboration: *When I work in groups in science class, there is team work.*	**6**	
First Language		Seeing Real World Connection: *In science class, I learn about the world outside of school.*	**6**	
Computer in the home		Belief that doing Well will help achieve Future Success: *Doing well in school won't help me have a satisfying career when I grow up.*	**6**	
Internet connectivity		Beliefs about Academic Press (that Teacher Presses them for Understanding): *My teacher asks me to explain how I get my answers.*	**5**	
Frequency of Museum visits		Use of Online Technologies Outside of School: *Outside of school, I use computers a lot.*	**3**	
		Attention and Perseverance: *Most of the time in science class I keep my mind on my work*	**3**	Responses to questions with a negative cast were reversed for scoring to give the equivalent positive rating.
		Organizational Skills:	**1**	

When I collect data, I try to use charts and other things to be organized

RESULTS

Demographics:

The demographics of the two testing sites are shown in Figures A1-A4 in Appendix B. Overall class sizes and gender groups were similar across the study, except for School A's control class. The control class in School A was smaller as it contained a group of students that were pulled out for ESL during science class instruction and thus were unavailable for this study (see Figure A1).

As was noted in the Procedure section, the two schools drew from different mixed-racial populations (see Figure A2). Indeed, as there were only four Caucasian students in the entire study, that variable was not included in the analysis. The variables African-American, Hispanic and Asian were included. The majority race was different in each school, with 64% of the students in School A identified as African-American, and 50% of School B identified as Asian (primarily Chinese).

22% of our students spoke Spanish as a first language, another 29% spoke Chinese (see Figure A3). A concern we had was whether the high English reading demand of MUVEES would discriminate against ESL students. To compensate for this, the project was originally designed with Spanish as an alternative. Unfortunately, since Spanish was not our only primary second language, this option was removed from the program for this implementation. However, at each level of our analysis, we carefully considered whether first language affected the results.

In three of our four classes (one class failed to report this data), 80% of our students were identified as being qualified for free or reduced lunch (see Figure A4). This designation was used in this project as a measure of low socio-economic status (SES), leading to an average of 80% of the students designated as low SES.

An intriguing result of this project can be seen in Figure 1, on the following page. The digital divide has been an oft-touted concern for many, and it has been generally accepted that computers in the home are more likely to be found in suburban middle class homes than in urban ones. Yet, despite the high percentage of disadvantaged urban students in this study, nearly all students, 93%, reported having a computer in the home with nearly 92% of them connected to the Internet. This supports new statistics on the prevalence of home computers (Russell, 2003).

Figure 1. Status of Technology at Students' Homes (n=81)

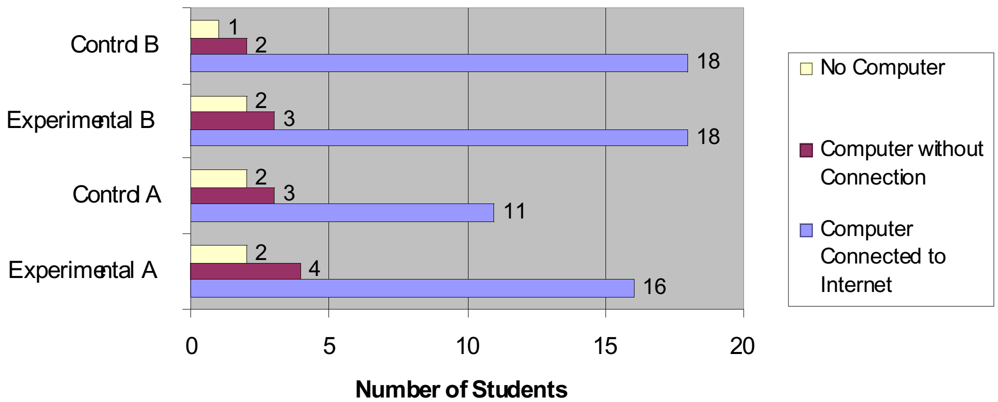
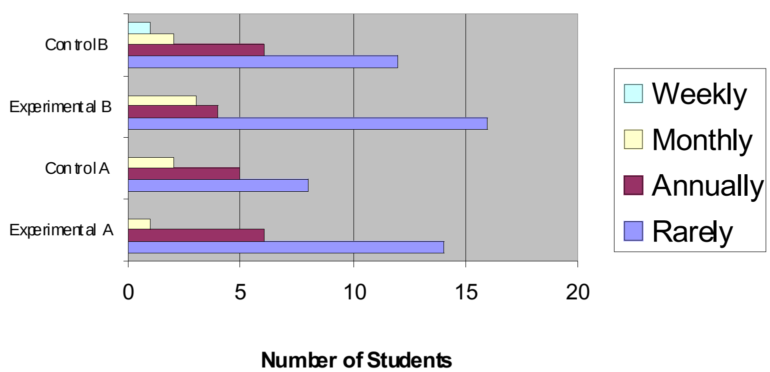


Figure 2 shows the frequency of museum visits as reported by students. The vast majority, 62.5%, report visiting museums only rarely, with 26.25%, visiting annually. One of the hoped-for goals from the Smithsonian participation in this project was to see if using the museum artifacts increased students' interest in and visits to museums. These students would appear to be a great target audience for this, and follow-up to discover if museum visits did increase would be recommended.

Figure 2. Frequency of Museum visits as Reported by Students (n=80)



Mean, standard deviation and ranges for the main variables in this study broken down by experimental and control group can be found in Tables A1 and A2 in Appendix C.
To summarize, the hypothetical "typical" student in the study is African-American, English-speaking with free/reduced lunch and a B- average in science this year. This hypothetical student owns a computer that is connected to the Internet and rarely visits museums.

Content learning

The content tests covered three main areas: scientific method literacy, knowledge of disease transmission and problem-solving skills. The overall raw pre- and post-scores for the content test show a 3% gain by the experimental group and a 4% gain by the control group—not a statistically significant difference, on average[2]. As can be seen from this chart, students, on average, started the project with 48% understanding of the Scientific Method as tested by this instrument, high for the start of a new unit. Scientific method is a topic that middle school students study throughout the year and throughout their schooling. This complicates looking at content gain on average. Had this been a new unit in which few if any students had prior knowledge, then the potential for gain is the entire range. In this situation, our potential for gain was limited to half the range, making average changes difficult to see in a 2-week intervention. Of more interest to this study was the answer to the research question of whether this intervention could help even the playing field for academically at-risk students.

Preliminary results indicate the MUVE is motivating and enhances learning for lower ability students. Six out of seven experimental students scoring less than 35% on the content pre-test improved their content knowledge above that level, while only two of five control students did so. We found that students did perceive different problems to solve in the MUVE's River City scenario. In our seventh grade classroom, e.g., there were five different hypotheses with causes ranging from population density to immigration to water pollution.

Table 3 shows the t- and *p* values for a regression model that predicts content posttest score. This model explains nearly 75% of the variation in the posttest scores.

Table 3. Final Model predicting Score on Content Posttest. (n=81)

Final Model	
Variable	**est B (se)**
Content Pretest Score	**0.835*** (0.088)**
Experiment	**-10.018* (4.50)**
Thoughtfulness of Inquiry at Start (TI)	**0.340 ~ (0.189)**

[2] The quantitative data was analyzed with SAS. A significance level of $p < .05$ was used and checks for linearity, normality and homoscedasticity were performed at various intervals. No clear violations were noted.

TI*Experiment (interaction term)			**0.263*** **(0.108)**
Science GPA			**5.450*** **(2.430)**
TI*ScienceGPA (interaction term)			**-0.154*** **(0.059)**
Hispanic			**-12.328*** **(5.166)**
Hispanic*TI (interaction term)			**0.300**** **(0.111)**
High SES[3]			**15.207**** **(4.947)**
HighSES*TI (interaction term)			**-0.367**** **(0.116)**
SUMMARY STATISTICS		$\mathbf{R^2}$	**74.45**
		F (df)	**13.7** **(10, 47)**
		***P* of F**	**.0001**

~ *p*<.10   **p*<.05   ***p*<.01   ****p*<.001

The regression equation is:

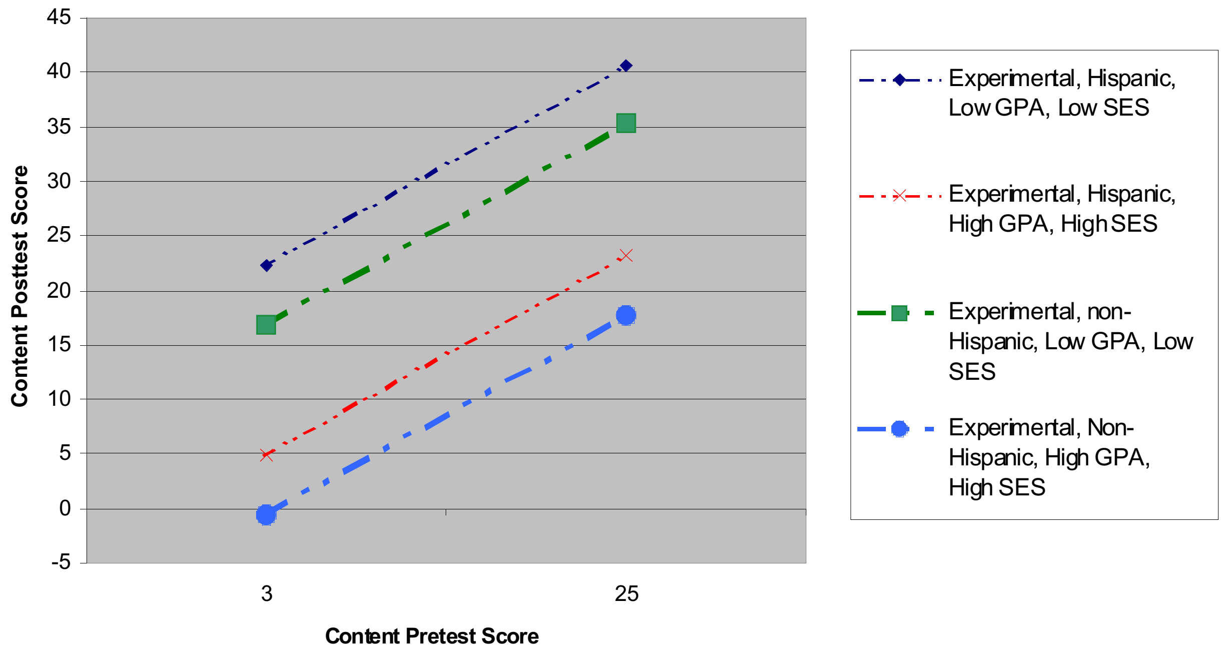
$$\text{Post test Score} = -7.48 + .84\text{Pretest} + -10.02\text{Experiment} + .34\text{TI} + .26\text{TI} * \text{Experiment} + 5.45\text{ScienceGPA} + -.15\text{TI} * \text{ScienceGPA} + -12.33\text{Hispanic} + .30\text{Hispanic} * \text{TI} + 15.21\text{HighSES} + -.37\text{HighSES} * \text{TI}$$

Figure 3 shows the relationship between content pre-test and post-test score for students with high perceptions of thoughtfulness of inquiry at the start of the project based on this model. As stated previously, this project presupposed that students had some pre-knowledge of scientific method and inquiry, so these students were the target audience[4]. As can be seen from this graph, at each content pre-test score, the experimental students with low Science GPA and on free or reduced lunch (low SES) scored the highest on the content post-test. Indeed, Hispanic students with low GPA and low SES improved their content score by approximately 20 points across the

[3] High SES is defined in this report as not being on free or reduced lunch
[4] For students with low perceptions of thoughtfulness of inquiry, results were reversed.

board! However, students with high GPA and high SES barely held even. While not shown, for every pre-test score (controlling for Science GPA and SES), the experimental students scored an average of approximately 5 points better than the control students did. Of additional interest was the fact that Hispanic students, regardless of native language, scored better then non-Hispanic students by nearly 5 points. This relationship, however, changes depending on students' initial metacognitive awareness of the inquiry process. This potentially indicates that there are students whose GPA is not reflecting their metacognitive awareness of inquiry, and that this MUVE has allowed these students to grow.

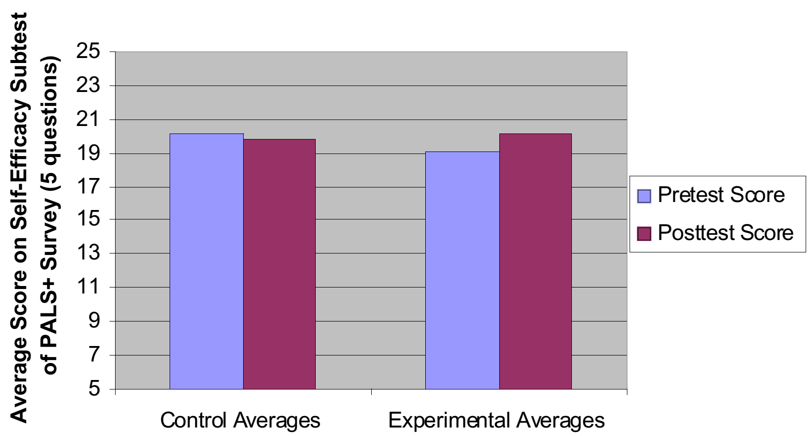
Figure 3. The Effect of Ethnicity, Science GPA and SES status on Content Posttest Score as a Function of Pretest Score for Students with High Perceptions of Thoughtfulness of inquiry at Start (n=81)



Efficacy

Academic efficacy (students' belief in their ability to master curricular knowledge and skills) addresses our question of improving learning for all students. Students who do not think they can successfully accomplish required educational tasks to gain desired learning outcomes will not try and, thus, will not learn. Improving academic self-efficacy, therefore, becomes important for students to have a chance at success. On average, students began the study with a good sense of overall academic efficacy. Both control and experimental groups averaged approximately just under 4 ("often true") on the five efficacy questions, on a scale of one to five.   Self-perceptions of academic efficacy for the experimental group increased by one point total over the five questions, on average, while the control group declined by .31 points (t=3.36, $p<.05$). Figure 4, on the following page, shows the average change for each group from pre to post testing.

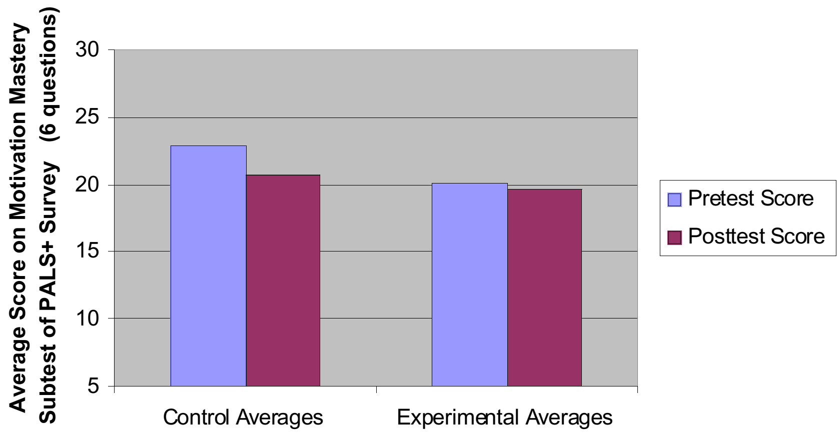
Figure 4. Average Academic Self-Efficacy Scores (n=76)



Motivation

The average change in self-perceptions of motivation over the six questions also showed significant differences between control and experimental group, (t=-2.1, $p<.05$) with the control group declining by an average of 2.4 points and the experimental group essentially unchanged (an insignificant decline of .4 points). Figure 5 shows these averages. While our hypothesis was that motivation would be improved absolutely through use of MUVEES, it is encouraging that we have a relative improvement over the control group. One might hypothesize that our novel presence in the classroom during the pretest might have caused such interest as to result in a higher than typical motivation, resulting in little change compared to the posttest when our presence was not an interesting novelty. In future implementations, to control for this possibility, the researchers should become an expected fixture in the classroom prior to administering the pretests or implementations.

Figure 5. Average Motivation Mastery Scores (n=76)



To help understand the effects of our intervention on motivation, a nested series of models[5] were developed[6] to discover predictors for changes in students' perceptions of motivation. No demographic variables were found to be statistically significant at the $p<.05$ level. Table 4 shows the variables and their *p* values.

Table 4. Final Model predicting for change on the Motivation Mastery of the PALS assessment. (n=81)

Final Model	
Variable	**est B (se)**
Experiment	**1.433~ (0.75)**
Science Interest	**0.202* (0.096)**
Collaboration	**0.128 (0.088)**

[5] The model was built by first adding the treatment variable (experiment[5]), then adding demographic variables and finally attitude variables.

[6] Checks for linearity, normality and homoscedasticity were performed at various intervals. No clear violations were noted

Seeing Connections to the Real World		**0.28**** **(0.100)**
Collaboration*ScienceInterest (interaction term)		**-0.046**** **(0.014)**
SUMMARY STATISTICS	**R^2**	**43.92**
	F (df)	**11.12 (5,71)**
	***P* of F**	**.0001**

~ *p*<.10   **p*<.05   ***p*<.01   ****p*<.001

The final equation predicting for change on the motivation section is shown below:

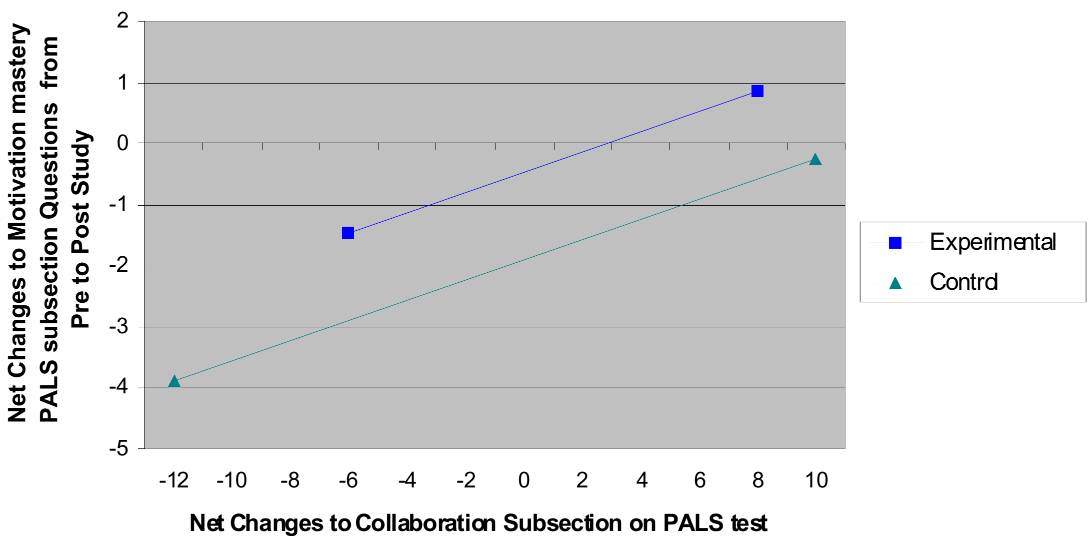
$$\text{Motivation} = -1.86 + 1.43\text{Experiment}^{6} + 0.20\text{ScienceInterest} + 0.13\text{Collaboration} + 0.28\text{RealWorld} + -0.05\text{Collaboration} * \text{ScienceInterest}$$

This equation indicates several things. First, there is a significant effect of the experimental group. On average, those students in the experimental group improved their motivation by 1.43 points more than those in the control group, controlling for collaboration, real world connections and science interest. In addition to this, for each four net points improvement overall for students response on the seeing real world connections subsection of the PALS test, there was associated approximately a one point increase in motivation on average, controlling for experimental group. While this is not surprising, given the design of our intervention and control curriculum, it is significant, as current demands to teach to standardized curriculum often results in loss of project-based units whose main goal is to connect science to the real world. As increasing motivation often results in increases in academic success, this finding lends support to the need to continue to offer real world curricula.

Figure 6, on the following page, shows the relationship between collaboration and motivation, controlling for intervention. At negative changes in collaboration, both groups show negative changes in motivation. Controlling for science interest and seeing real world connections, the experimental group is associated with less negative changes in their motivation mastery than are control students. Indeed, the control students are associated with declining perceptions of motivation mastery at all levels of collaboration, whereas, at more positive changes in collaboration, experimental students show increases in motivation mastery. This might indicate that, since the program requires students to work collaboratively for success, those that were able to experience growth in this area were able to assimilate more of the learning embedded in the program and thus experience positive changes in motivation.

This model has an R^2 of nearly 44%, indicating that a considerable amount of variation in net motivation scores still needs to be explained. It is possible that the relationships discussed above will not persist once different predictors are identified.

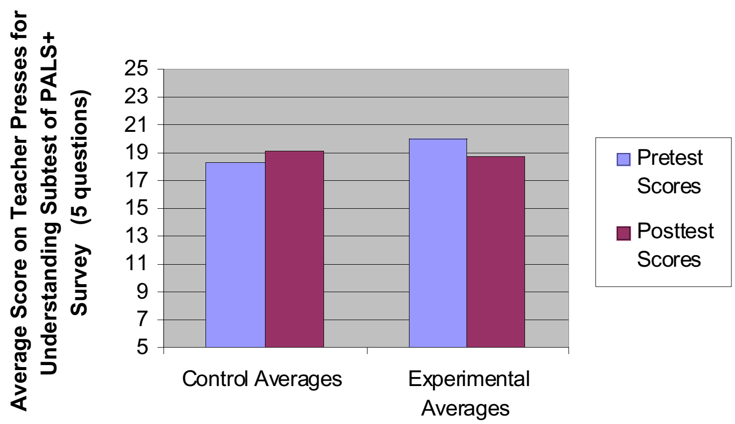
Figure 6. The Effect of Changes in Collaboration over the length of the Study on Net Changes in Motivation Mastery subsection of the PALS+ Survey, controlling for Treatment Group (real world and science interest set to mean) n=81



Academic Press

Of interest are our findings about academic press. This variable measures the student's perception of how much their teacher presses them for understanding. There is a significant difference in the means between the experimental and control groups (t=-2.64, $p<.05$), with the control group beliefs increasing by .41 and the experimental group declining by 0.93, on average (Figure 7, next page). While other studies have shown a positive correlation between academic efficacy and academic press (Middleton, 2002), the effect of academic press is opposite that of academic efficacy in this study.

Figure 7. Average Perceptions of Teacher Presses Me for Understanding



At first glance, this result appears to indicate that students were not asked to reach understanding. However, the emphasis in these questions is on the students' perceptions of how much the teacher *pressed* the students (see Table 2). It is possible that the decline in teacher press for the experimental group reflects the more autonomous nature of the experimental group, as opposed to indicating that they are not being required to reach understanding. This might indicate the role technology could play in supporting students' growth towards self-responsibility in learning. More research into this and effects of the intervention on students' understanding are necessary to tease out the underlying explanation for this difference.

Change in Multiple Choice Section of the Content Test

Another series of nested regression models were run for the outcome of "change in the Multiple-choice section of the content test." This was chosen for theoretical reasons. The most common complaint about technology in education is that the results are not reproducible on standardized multiple-choice paper tests, which are major criteria for most policy-makers. The change in the results from the multiple-choice part of the content test from pretest to posttest were segregated from the results as a whole and used to create a new variable. A nested series of models were developed[7]. Table 5 on the following page shows the variables and their *p* values. This model only explains approximately 37% of the variation in the multiple-choice section, leaving much variation unexplained. This variation could be due to individual variation, testing process, or unidentified predictors.

---

[7] Checks for linearity, normality and homoscedasticity were performed at various intervals. No clear violations were noted

Table 5. Final Model predicting for change on the Multiple Choice Section of the Content Test. (n=81)

Final Model		
Variable		**est B (se)**
Content Pretest Score[8]		**-0.093~ (0.048)**
Experiment		**-0.113 (0.54)**
Attention		**-0.279** (0.095)**
Museum Visit Frequency		**0.644** (0.282)**
Technology Interest		**0.142* (0.069)**
Use of Technology outside School		**-0.145 (0.148)**
Technology outside*Experiment (interaction term)		**0.481* (0.195)**
Grade [9]		**-0.354 (0.654)**
Grade*Experiment (interaction term)		**2.026* (0.848)**
SUMMARY STATISTICS	**R^2**	**36.73**
	F (df)	**4.19 (9, 65)**
	***P* of F**	**.0003**

~ *p*<.10   **p*<.05   ***p*<.01   ****p*<.001

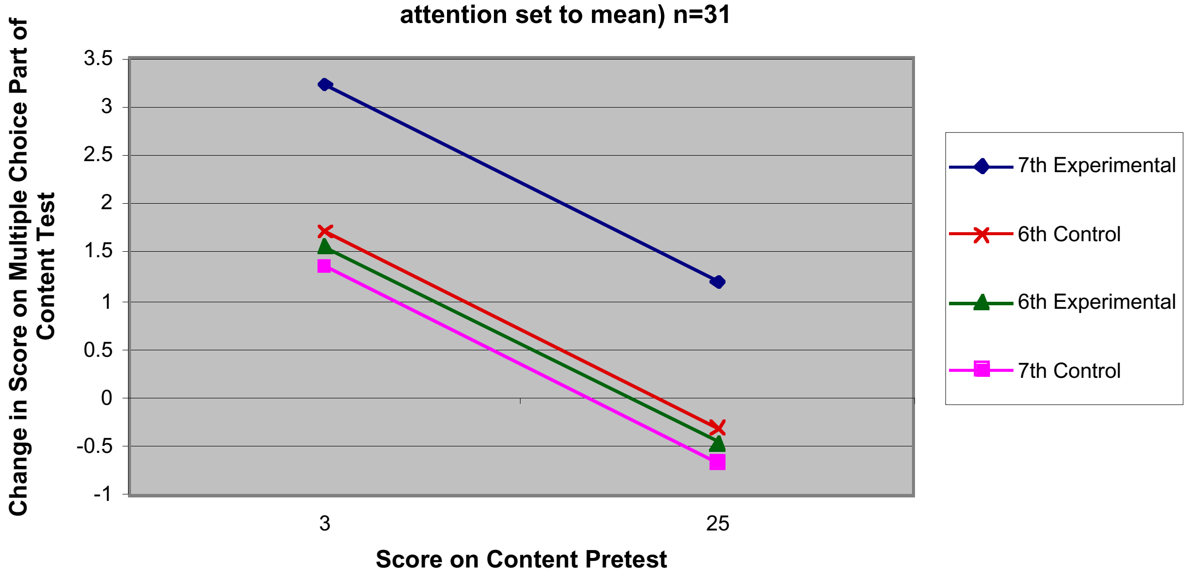
[8] Content Pretest score is significant at p<.07 level and was included in this study due to its proximity to the arbitrarily chosen standard of p<.05 level.
[9] Grade is not significant when it equals one (7th grade) and experiment =0 (Control), thus, the following analysis does not include the control class for 7th grade

On average, there is a positive association between students' perception of their home technology (video games, computer, chat) use and their change in score on the multiple-choice section of the content test. The subset of students in the experimental group who perceived their computer use increasing during the study had positive changes in the multiple choice section of the content test. However, students in the control group had the reverse effect. It is tempting to wonder if use of the project software prompted more constructive use of technologies at home, thus improving learning overall. Of course, it is equally as possible that students in the control group changed their home game-playing due to envy of their peers in the experimental group! More investigation into this phenomenon would be needed to understand it.

Controlling for home computer use, students in the experimental group with low pretest scores had two points higher improvements in their multiple choice scores than did experimental students with high pretest scores. While it is not surprising that students with high scores did not improve greatly, as they had less room to do so, it is encouraging that those students with low test scores improved significantly. One of the research questions this study hoped to answer is if emerging technologies could 'level the playing field' for disenfranchised students. This result gives some credence to that idea that lower achieving students can make up ground using MUVEs.

Figure 8, following page, shows a large gain for the 7^{th} grade experimental group over all of the other classes. While the 7^{th} grade control class is not inferable to the population, it is included here for discussion of this sample. This figure shows that 7^{th} grade experimental students (regardless of content pretest score) were associated with positive changes in their multiple-choice score. Additionally, these students scored at least 1.5 points higher than the other three classes. Potentially, this indicates that this technology is better suited to the older middle school student. It would be interesting to compare middle-schoolers with high-schoolers to see if the effect persists.

Figure 8. The Relationship between Content Pretest Score and Changes in Multiple Choice Questions over the Course of the Study, controlling for Grade and Treatment(Tech Interest, Outside Use of Technology, Freq museum visits, attention set to mean) n=31



SUMMARY OF ANALYSIS OF MUVE BOSTON IMPLEMENTATION

Content:

- The greatest improvement on the multiple-choice section of the content test was for the experimental groups with low pretest scores. This is important, as it indicates that we can improve learning for our lowest students as measured by standardized tests while using engaging pedagogical strategies.

- Interestingly, experimental students who perceived themselves as using more technology at home by the end of the project than at the start had positive changes in their multiple choice scores. This potentially indicates that using technology in the school leads students to using it more effectively at home.

- Experimental students with low science GPA and low SES had the highest content posttest score of any students, if they started with high metacognitive awareness of the inquiry process.

- Experimental students scored five points higher than control group students at all levels of pretest scores, if they entered with high perceptions of their own thoughtfulness of inquiry.

Academic Efficacy and Press:

- On average, the experimental students increased their rating of their academic efficacy while the control students decreased theirs slightly.
- Over the course of the study, the control group increased their view that their teacher pressed them for understanding, while the experimental group decreased theirs. This may reflect the new role for the teacher that she took in the experimental group as a facilitator, not the leader, and possibly explains the increase in the academic efficacy that experimental students achieved. As most schools have a long-term goal of creating self-learners, this is a positive outcome of the MUVE as a tool to support students' growth towards self-responsibility in learning.

Motivation:

- On average, students in the experimental group improved their motivation as compared to the control group.
- There is a relationship between students' perceptions of improved collaboration skills and increased motivation.

Grade:

- 7^{th} grade experimental students showed approximately five points more improvement at all levels of content pretest scores than did 6^{th} grade students. This might be a reflection on the curriculum chosen, the demands of this unit or even the different approaches taken by each teacher. More research would be needed to explore this.

Race/Gender:

- Hispanic students with high metacognitive awareness of the inquiry process scored better on the content test than non-Hispanic students.

Technology:

- Interestingly, despite our students' inner-city designation, 93% of them reported having a computer at home with 92% of those connected to the Internet.

Miscellaneous:

- The variable denoting native and non-native English speaking students was not significant in any of the 4 regression models built, despite the fact that over 50% of the students were designated ESL (see Figure A3, Appendix B). This is significant, as it indicates that language was not a barrier to success in using the MUVE.

DISCUSSION

This project was initiated to explore the affordances of MUVEs in middle school science. The process cycled between design, implementation, evaluation and redesign. In each iteration, we discovered methods to improve the implementation process and learning outcomes.

Despite a major emphasis on developing insights about how to design MUVEs for learning, we were able to make several discoveries about their potential in the classroom. MUVEs appear to be motivating and to improve academic self-efficacy. In addition, MUVEs seem to have a greater impact on learning for the bottom third of students. Three girls from the 7th grade experimental classroom offer a case in point. These girls were intrigued with our initial presence in the classroom; but, by the time the pre-testing was concluded, they had reverted to their previous inattentive behaviors. In fact, for the first day or two of the intervention, they groaned when we entered the room. Two weeks later, these girls were still inattentive to their teacher; but, instead of passing notes on non-academic issues, they were working hard on finishing the project and demanded that I read their final report immediately. While anecdotal, this example was not isolated and potentially indicates a role for MUVEs in the classroom.

We have also learned much on the technical strengths and weaknesses of MUVEs in the classroom. In order to facilitate chat, computers needed to be connected to the Internet. This sparked concerns from technology administrators and left the project vulnerable to bandwidth and filtering issues. Several iterations of MUVEES had to be tried before a smoothly working one was found.

Due to the novelty of this technology, professional development for teachers in the classroom needs to be intensive. The teachers involved in the first implementation were part of the design team and thus understood the purposes and set-up of MUVEES. The teachers in the Boston implementation were a better example of the issues in training needed for MUVEs such as MUVEES to scale up. While they were given several hours of professional development prior to the start of the project, and were also given an in-depth teachers manual, they both felt that additional training would have been advantageous. Interestingly, while we had hoped that they would each integrate MUVEES into their own curriculum, one teacher felt confined by the software and the other wanted it to be more proscribed, despite apparently similar philosophical approaches to teaching.

CONCLUSIONS AND AREAS OF FUTURE RESEARCH

MUVEs are an emerging technology that has already captured the attention of game designers and players. This report represents an initial attempt to develop heuristics for their use in the K-12 classroom. The data is promising, but not conclusive about the effectiveness of MUVEs for learning. By examining student interactions with the pilot curriculum, we saw ways to strengthen our content and pedagogy. We also saw teachers struggle with facilitating the whole-class interpretive sessions that alternated with MUVE experiences—an indication that we need to extend our professional development experiences. Some possible areas of future research not addressed by this study might be to: assess the effect of integrating museum content on students' museum visits, investigate learning gains in a new content area across a range of grades, and create bonuses in-world for learning content knowledge.

IMPORTANCE OF THIS STUDY

Our project is extending current MUVE capabilities in order to study the science learning potential of immersive simulations, interactive virtual museum exhibits and "participatory" historical situations, particularly for low-performing students who are turned off to school and skeptical about their ability to learn. Our research to date demonstrates that MUVEs seem quite feasible as an addition to more conventional kinds of computer-based instruction. Our long-range goal is to develop "replacement" units that—through simultaneously developing students' knowledge in science, mathematics, history, and social studies—enable the teaching staff to utilize inquiry methods capable of covering the same span of content in a similar amount of time required by presentational teaching of this material one topic at a time. Given current educational "reforms" based on an overcrowded curriculum driven by high-stakes, low-level tests, this is essential for any scalable curricular intervention.

REFERENCES

Dede, C., Ketelhut, D., & Ruess, K. (2002, October 23-26, 2002). *Motivation, Usability, and Learning Outcomes in a Prototype Museum-based Multi-User Virtual Environment.* Paper presented at the ICLS, Seattle, Washington.

Middleton, M. J. (2002). *Do the effects of classroom demands depend on student motivation? The interaction of academic press and achievement goals*, [Internet]. Available:

http://edtech.connect.msu.edu/searchaera2002/viewproposaltext.asp?propID=3564March 9, 2003].

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.

Russell, M. (2003). Educational Technology and Leadership, Preliminary Findings: Guest Lecture. Cambridge, Ma.

Tobin, M. (1999). Improving Student Retention Through the use of Technology. *Unpublished Masters*, Saint Xavier University.

This material is based upon work supported by the National Science Foundation under Grant No.REC-0202543. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

ILLUSTRATIONS

Figure 1: Birds-eye View of River City

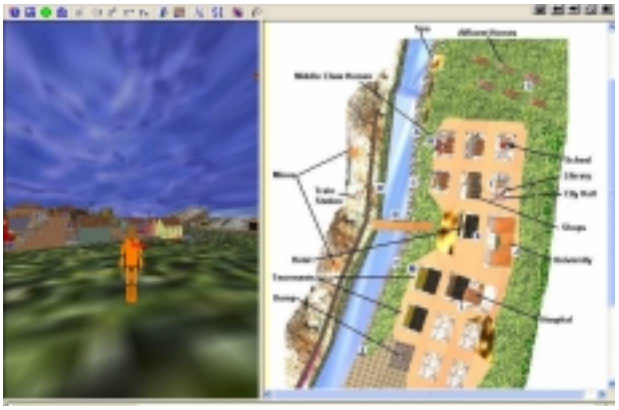


Figure 2: A Street Scene in River City

Figure 3: Lab Equipment inside a Building

Figure 4: River water sampling station

Figure 5: Screen Shot of a Non-Player Character (NPC) agent giving information.

Appendix A: Survey Instruments

SCIENCE AND TECHNOLOGY ATTITUDES AND BELIEFS

Last name: __________________________

First name: __________________________

School: _____________________________

Teacher: ____________________________

Science Period:

Your gender

O Male

O Female

For office use only

ID number

_
_
_
_
_
_
_
_
_
_

_
_
_
_
_
_
_
_
_

EXAMPLE QUESTION:

	1 not true at all		3 somewhat true		5 very true
I like strawberry ice cream.	_	_	_	_	_

Please answer the following questions about your experience in science class this year. Indicate your answer by completely filling in one of the numbered circles — from Number 1 (not true at all) through Number 5 (very true). You can use any number on the scale.

	1 not true at all		**3 somewhat true**		**5 very true**
1. An important reason why I do my work in science is because I like to learn new things	_	_	_	_	_
2. In science class I am asked to explain the results of investigations	_	_	_	_	_
3. When doing my work in science, I stop once in a while and go over what I have done	_	_	_	_	_
4. In science class I find it easy to concentrate on what we are learning	_	_	_	_	_
5. I like learning to use technology	_	_	_	_	_
6. I try to figure out the main ideas of what I am learning about in science class	_	_	_	_	_
7. I'm certain I can master the skills taught in class this year.	_	_	_	_	_
8. I think learning about science is useful	_	_	_	_	_
9. In science class I am asked to think about the evidence for statements	_	_	_	_	_
10. In science class the teacher demands that I do thoughtful work	_	_	_	_	_
11. Learning science can help me understand the differences between people	_	_	_	_	_

TURN THE PAGE

	1 not true at all		3 somewhat true		5 very true
12. My science classroom is a fun place to be	_	_	_	_	_
13. I'm certain I can figure out how to do the most difficult class work	_	_	_	_	_
14. I like learning about science	_	_	_	_	_
15. To help me understand things I do for science, I make outlines, concept maps or pictures	_	_	_	_	_
16. I can do almost all the work in class if I don't give up.	_	_	_	_	_
17. Even if the work is hard, I can learn it.	_	_	_	_	_
18. I like learning science even if I make a lot of mistakes	_	_	_	_	_
19. The ideas we learn about in science class are related to what s happening in our city	_	_	_	_	_
20. I can do even the hardest work in this class if I try.	_	_	_	_	_
21. Even if I do well in school, it will not help me have the kind of life I want when I grow up.	_	_	_	_	_
22. When I work in groups in science class, there is teamwork	_	_	_	_	_
23. When work is hard in science, I either give up or do only the easy parts	_	_	_	_	_
24. I do my science work because I m interested in it	_	_	_	_	_

TURN THE PAGE

	1 not true at all		3 somewhat true		5 very true
25. In science class I explain my ideas to other students	_	_	_	_	_
26. In science class, I learn about how to make things better in life.	_	_	_	_	_
27. I learn from other students in science class	_	_	_	_	_
28. Students ideas and suggestions are used during science class	_	_	_	_	_
29. What we do in science class is a waste of time	_	_	_	_	_
30. My chances of succeeding later in life don t depend on doing well in school.	_	_	_	_	_
31. Doing well in school doesn t improve my chances of having a good life when I grow up.	_	_	_	_	_
32. Most of the time in science class I keep my mind on my work	_	_	_	_	_
33. In science class I find out answers to questions by doing investigations	_	_	_	_	_
34. Understanding how to use technology is important to me	_	_	_	_	_
35. In science class the teacher accepts nothing less than my full effort	_	_	_	_	_

TURN THE PAGE

	1 not true at all		3 somewhat true		5 very true
36. When I answer a question in science, the teacher often asks me to explain why I think it s the right answer	_	_	_	_	_
37. In science class, I learn about the world outside of school	_	_	_	_	_
38. I like science work best when it really makes me think	_	_	_	_	_
39. In science class, I don't always pay attention to what is being said	_	_	_	_	_
40. I think learning about technology is useful	_	_	_	_	_
41. When I collect data, I try to use charts and other things to be organized	_	_	_	_	_
42. Getting good grades in school won t guarantee that I will get a good job when I grow up.	_	_	_	_	_
43. Outside of school, I use computers a lot.	_	_	_	_	_
44. In science class students cooperate with each other to learn new ideas	_	_	_	_	_
45. Understanding science is important to me	_	_	_	_	_
46. I often chat on-line with my friends after school.	_	_	_	_	_

TURN THE PAGE

	1 not true at all		3 somewhat true		5 very true
47. When I design investigations, I try to think about how I can get the most information	_	_	_	_	_
48. Even if I am successful in school, it won t help me fulfill my dreams.	_	_	_	_	_
49. I often visit museums.	_	_	_	_	_
50. In science class, the teacher doesn t allow me to get away with doing easy work	_	_	_	_	_
51. When I look at my data, I try to think about how it answers the question	_	_	_	_	_
52. In science class, I learn how science can be part of my out-of-school life	_	_	_	_	_
53. When I am learning science, I try to make the ideas fit together	_	_	_	_	_
54. I think learning about technology is a waste of time	_	_	_	_	_
55. Doing well in school won t help me have a satisfying career when I grow up.	_	_	_	_	_
56. When I prepare presentations, I check to be sure that I include the main ideas of the project	_	_	_	_	_

TURN THE PAGE

	1 not true at all		3 somewhat true		5 very true
57. I enjoy playing computer games.	_	_	_	_	_
58. In science class, the teacher doesn t just accept my answer but wants me to show I understand	_	_	_	_	_
59. I would rather be in my science class than any other one	_	_	_	_	_
60. I try to connect the things I am learning about in science with what I already know	_	_	_	_	_
61. I think learning about science is boring	_	_	_	_	_
62. Thinking about the real world helps me learn in science class	_	_	_	_	_
63. An important reason why I do my science work is because I want to get better at it	_	_	_	_	_
64. I enjoy what we do in science class	_	_	_	_	_
65. In science class, students question each other s ideas in order to figure things out	_	_	_	_	_

Thank you for sharing your thoughts with us.

We appreciate your time and attention!

MUVEES CONTENT SURVEY

Last name: ________________________

First name: ________________________

School: ___________________________

Teacher: __________________________

Science Period: _____________________

Circle the best answer below:

1. Your gender: male female

2. Language spoken at home: English Spanish Haitian Other: ___________

3. Do you visit museums: one time/week one time/month one time/year rarely

4. Have you ever been to the National Museum of American History in Washington?
   Yes No

5. Do you or your family own a computer? Yes No

Scientific Method Survey Questions

Section A: The following questions, #1-15, are multiple choice. Please circle the letter of the ONE response in each item that best answers the question or completes the statement.

1. When using the scientific method, the first step is to ___________
   a. Gather information on the problem
   b. Form a hypothesis
   c. State the problem
   d. Record and analyze data
2. What is the last step of the scientific method?
   a. Experimenting
   b. Stating a conclusion
   c. Forming the hypothesis
   d. Stating the problem
3. Which phrase describes a hypothesis?
   a. Reviewing information related to a problem
   b. Testing a factor, or variable, in an experiment
   c. Suggesting a possible solution to a problem after studying it carefully
   d. Performing a controlled experiment
4. The discovery that bacteria caused specific diseases happened in the time period
   a. 1800-1850
   b. 1850-1900
   c. 1900-1950
   d. after 1950

5. A hypothesis is formed _______________

    a. Before the problem is stated

    b. Before the experiment is done

    c. After the conclusion is stated

    d. After analysis of the data

6. The system that scientists use to solve problems is called ____________

    a. Experimental treatment

    b. Data collection

    c. Scientific method

    d. Adaptive response

7. When you write down step by step how you are going to do your experiment, you are writing a (an) _________

    a. Procedure

    b. Data

    c. Conclusion

    d. Objective

8. Diseases can be spread by eating contaminated food and by ______________

    a. Going outside without your coat on

    b. Insects

    c. Books

    d. None of the above

9. ___________ are what a person performing an activity sees, hears, smells, feels or tastes.

   a. Observations

   b. Theory

   c. Variable

   d. Inference

10. You are conducting an experiment to determine if putting an additive in gasoline will improve gas mileage. All cars used are identical. The gasoline used in each car is the same. The car used as the experimental <u>control</u> contains _____________

   a. One part gasoline and one part additive

   b. Two parts gasoline and one part additive

   c. Only additive

   d. Only gasoline

11. If you were testing how well different laundry soaps cleaned your clothes, the <u>independent variable</u> would be the _____________

   a. How clean the clothes are

   b. Soap

   c. Water

   d. Washing machine

12. Which of the following conditions would might **<u>increase</u>** the likelihood that a contagious disease would be spread:

   A. living with a sick person
   B. living in a hot, humid place
   C. being wealthy

   a. A only

   b. A and B

   c. C only

   d. A, B and C

13. If you were testing how well different laundry soaps cleaned your clothes, the <u>dependent variable</u> would be the **_____________**

    a. How clean the clothes are

    b. Soap

    c. Water

    d. Washing machine

14. Flush toilets started to be used in private homes in what time period?

    a. 1800-1850

    b. 1850-1900

    c. 1900-1950

    d. after 1950

15. Microbe is another name for a **_________**

    a. Bacteria

    b. Plant

    c. Microscope

    d. Special dress

Section B: The following questions, #16-18, are about the story below. Please read the problem and then write a statement that <u>best</u> answers the question or completes the statement.

After many observations, you find that your bicycle tires look flatter on cold winter days than they do on hot summer days---even though you fill them with the same amount of air.

16. State the <u>problem.</u>

17. Form a possible <u>hypothesis.</u>

18. How would you test this hypothesis?

Section C: Read the following statements, labeled A-H, below. Please write the letter of the statement that best answers the question or completes the statement in questions # 19-24 .

A. A scientist wants to find out why seawater freezes at a lower temperature than fresh water.

B. The scientist goes to the library and reads a number of articles about the physical properties of solutions.

C. The scientist also reads about the composition of seawater.

D. The scientist travels to a beach and observes the conditions there.

E. After considering all of the information, the scientist sits at a desk and writes, My guess is that sea water freezes at a lower temperature than fresh water because sea water has salt in it

F. The scientist goes back to the lab and does the following:

1. **Fills two beakers with 1 liter of fresh water**
2. **Dissolves 35 grams of salt in one of the beakers**
3. **Places both beakers in a freezer**
4. **Leaves the beaker in the freezer for 24 hours**

G. After 24 hours, the scientist examines both beakers and finds the fresh water to be frozen. The salt water is liquid.

H. The scientist writes in a notebook.

19. Which statement contains **procedure steps?**

20. Which statements contain **observations?**

21. Which statement contains the **hypothesis?**

22. What is the **independent variable?**

23. Which beaker is the **control?**

24. Finish **Statement H** above by writing a conclusion for the scientist.

Section D: Read the following story. Please answer questions #25-26 that follow based on your interpretation of this story.

You have been hired to help solve a problem for a local middle school basketball coach. In your investigations, you discover the following: There are two teams, one in the country whose coach hired you, and one in the nearby city.

The country team practices at their school for 1 hour every day before the students return to their farms, traveling about 45 minutes to get home. Their coach has been coaching at that school for about 20 years, and has had winning teams for the last 10 years.

The city team also practices at their school, in a brand new gymnasium, for 1 hour every day before the students return to their homes. Most of the students live in neighboring apartment buildings, only traveling about 10 minutes to get there. Their coach is only in her 3rd year, and this is her first winning season.

These two teams play 4 games together every season. Their 4th match is coming up, and the country coach who hired you wants to know why he has lost the first 3 games this year to the city coach, when previously his team won 3 out of 4 games each season.

25. Please state a **hypothesis** to present to the country coach about why you think he lost 3 games.

26. Outline, briefly, the steps you would take to test your hypothesis in the space below.

Section E: Please answer the following question in the space provided.

27. List the steps of the scientific method in order in the space below.

Thank you for sharing your thoughts with us.

We appreciate your time and attention!

Appendix B: Descriptive Statistics

Figure A1. Class Sizes broken down by School and Treatment (n=81)

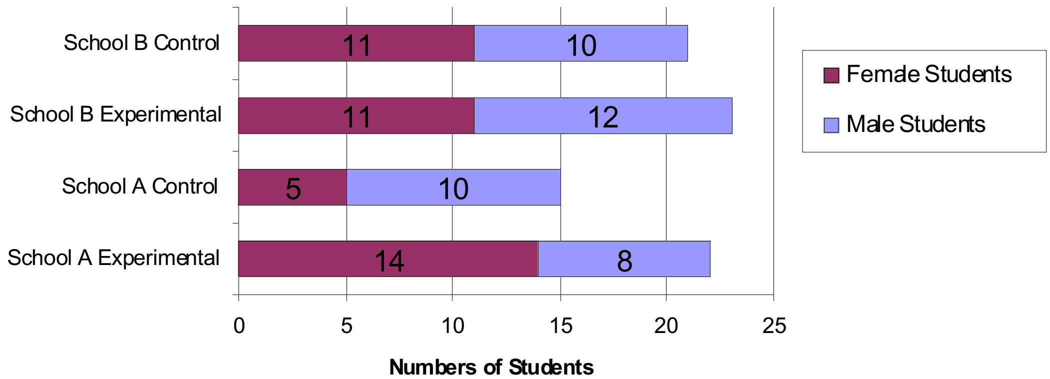


Figure A2. Racial Make-up of Each Class, broken down by School and Treatment (n=81)

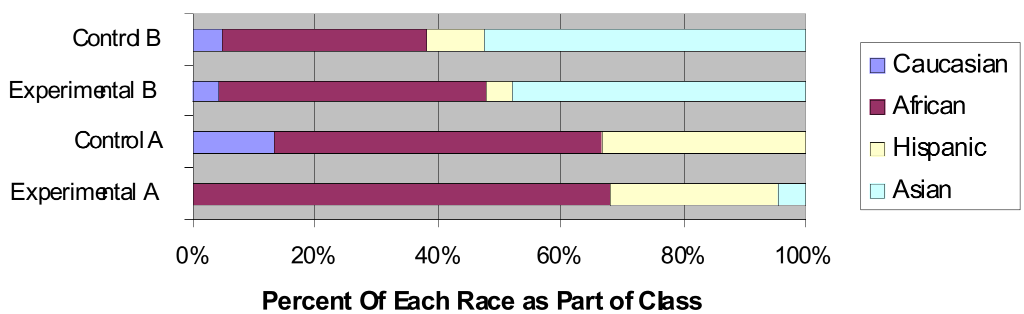


Figure A3. Primary Language Spoken at Home (n=81)

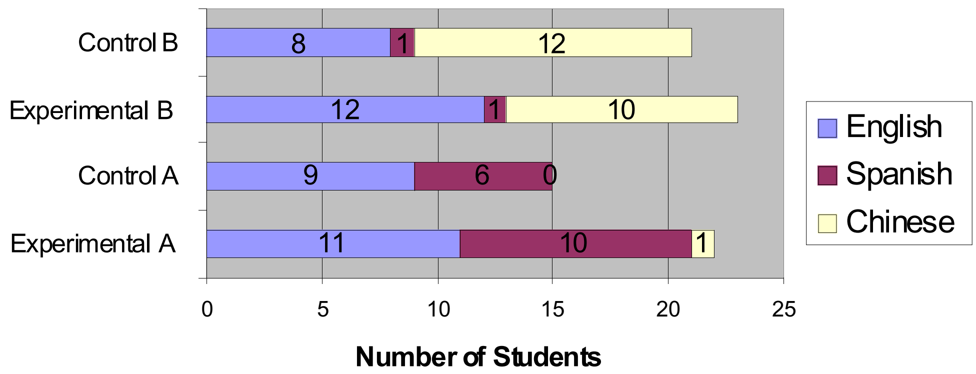
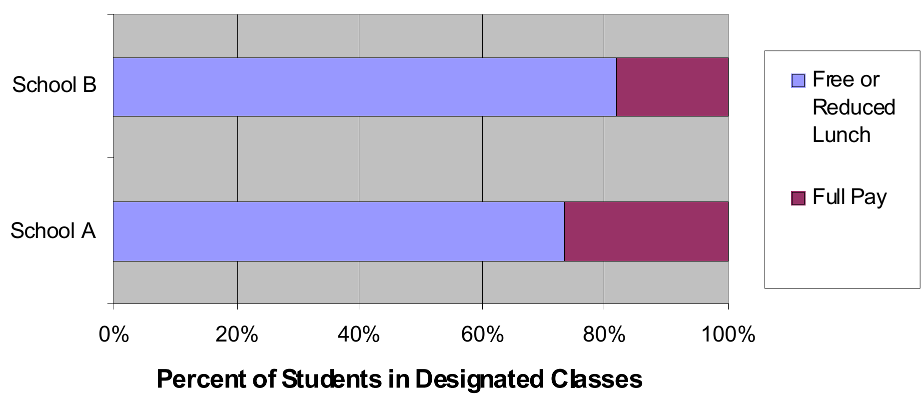


Figure A4. Status of Free or Reduced Lunches



Appendix C: Descriptive Statistics

Table A1. Simple Statistics for the Major Variables Collected in Study (n=81)

	Experimental (n=45) or Control (n=36)	**Mean**	**Standard Deviation**	**Low**	**High**	**Total possible**
Teacher Expectation of Success	***E***	***3.23***	***1.16***	***1***	***5***	***5.0***
	C	3.86	.86	1.5	5	
Science GPA Achieved this year	***E***	***2.68***	***1.02***	***0***	***4***	***4.0***
	C	2.69	.81	1	4.1	
Score on Content Pretest	***E***	***15.51***	***5.08***	***3***	***25***	***33***
	C	16.17	4.85	5	25	
Score on Content Posttest	***E***	***16.44***	***4.41***	***6***	***26***	***33***
	C	17.56	4.83	7	26	

Table A2. Simple Statistics for the Change in PALS+ Subsections from Pretest to Posttest (n=81)

	Experimental (n=45) or Control (n=36)	Mean	Standard Deviation	Low	High	Range possible
Change in Academic Efficacy Rating [10]	***E***	***1.07***	***2.71***	***-6***	***6***	***-20:20***
	C	-0.31	3.88	-11	7	
Change in Motivation Mastery[10]	***E***	***-0.49***	***3.57***	***-11***	***8***	***-25:25***
	C	-2.41	4.47	-15	4	
Change in Science Interest [10]	***E***	***-0.95***	***3.64***	***-8***	***6***	***-25:25***
	C	-0.44	5.10	-11	13	
Change in Thoughtfulness of Inquiry[10]	***E***	***0.78***	***6.79***	***-18***	***12***	***-55:55***
	C	.94	8.77	-14	31	
Change in Technology Interest[10]	***E***	***-1.34***	***3.11***	***-12***	***4***	***-15:15***
	C	-0.97	3.20	-11	6	
Change in Collaboration[10]	***E***	***1.11***	***4.09***	***-6***	***8***	***-25:25***
	C	-0.09	5.18	-12	10	
Change in Seeing Real World Connection[10]	***E***	***-0.16***	***4.18***	***-9***	***11***	***-25:25***
	C	1.0	4.87	-11	14	
Change in Belief that doing Well will help them achieve Future Success[10]	***E***	***0.95***	***6.00***	***-12***	***16***	***-25:25***
	C	.35	4.73	-10	11	
Change in Student Beliefs about Academic Press[10]	***E***	***-0.93***	***3.33***	***-10***	***6***	***-20:20***
	C	0.41	3.34	-8	8	
Change in Use of Online Technologies Outside of School[10]	***E***	***-0.23***	***2.29***	***-8***	***6***	***-10:10***
	C	0.09	2.11	-6	6	
Change in Attention and Perseverance[10]	***E***	***0.23***	***1.77***	***-4***	***4***	***-10:10***
	C	0.03	2.63	-5	7	
Change in Organizational Skills[10]	***E***	***0.29***	***1.16***	***-2***	***3***	***-4:4***
	C	0.28	1.23	-3	3	

[10] Change is calculated by subtracting the pretest score on the PALS inventory sub test from the post score on the same test at end of study.