Computer Supported Collaborative Learning
DOI 10.1007/s11412-008-9049-1

Exploring embedded guidance and self-efficacy in educational multi-user virtual environments

Brian C. Nelson · Diane Jass Ketelhut

Received: 8 September 2007 / Accepted: 11 August 2008

Abstract In this paper, we present the results of an exploratory study into the relationship between student self-efficacy and guidance use in a Multi-User Virtual Environment (MUVE) science curriculum project. We describe findings from a sample of middle school science students on the combined impact on learning of student self-efficacy in scientific inquiry and use of individualized guidance messages, and on the interplay between levels of self-efficacy and use of an embedded guidance system in an educational MUVE. Results from our study showed that embedded guidance was associated with improved learning outcomes for learners across a spectrum of self-reported efficacy in science. However, we also found that learners with low levels of initial self-efficacy in science viewed fewer guidance messages than their higher efficacy peers, and did not perform as well as their higher efficacy peers regardless of guidance use level. At the same time, outcomes for low self-efficacy students who used the guidance system heavily were raised to the level of high self-efficacy students who did not use the system.

Keywords Guidance · Inquiry · MUVEs · Self-efficacy · Science

Introduction

The once fantastical idea of embedding curricula into the game-like cyber worlds of multi-user virtual environments (MUVEs) is beginning to gain wider support in the educational research community. Increasing numbers of studies are investigating the use of these immersive computer-based environments as platforms for authentic inquiry in science

B. C. Nelson (✉)
Arizona State University, P.O. Box 870611, Tempe, AZ 85287, USA
e-mail: Brian.Nelson@asu.edu

D. J. Ketelhut
Temple University, Philadelphia, PA, USA

(e.g. Barab et al. 2005a; Nelson 2007; Nelson et al. 2007; Clarke and Dede 2005), and as environments supportive of situated learning principles and embedded activities that approximate authentic science (e.g. Gee 2003; Steinkuehler 2004; Steinkuehler and Chmiel 2006; Shaffer 2006; Galarneau and Zibit 2008).

Research in educational MUVEs is part of a broader focus on the use of gaming environments for learning (Gee 2003). For example, in 2006 the Federation of American Scientists released a major report in which they urged increased research into and financial support for investigating the use of complex gaming environments as platforms for learning. In the same year, the MacArthur foundation launched a five-year, $50 million “Digital Media and Learning” initiative to understand how digital technologies are shaping and changing the lives of young people. A principal component of the initiative supports the development and study of collaborative gaming environments including MUVEs.

Since 2002, we have been investigating the viability and learning impact of one such MUVE called River City. We are focusing on a variety of theory-based questions about use of the environment and associated curriculum to support instruction in middle school science inquiry. Our design definition for scientific inquiry stems from the National Science Education Standards:

> Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results (National Research Council 1996, p 23).

In the current study, we examine the possible interaction between student-perceived self-efficacy as practitioners of science and the use of an embedded guidance system which provides meta-cognitive questions. These questions are designed to support learners in reflecting on their inquiry and data gathering processes in a MUVE-based scientific inquiry curriculum. We are interested in the impact of self-efficacy since the literature in this area indicates that it mediates behavior (see for example, Pajares 2000) and our own research demonstrates that scientific inquiry behaviors are impacted by a student’s perceived self-efficacy (Ketelhut 2007). Research in psychology and education also tells us that we might expect the level of a student’s self-efficacy in scientific inquiry to impact the use of learning strategies, such as our guidance system (Pintrich and DeGroot 1990). From this we hypothesize that increased levels of self-efficacy are associated with increased levels of guidance viewings.

To further complicate this relationship, both levels of self-efficacy and of guidance are known to positively influence performance on tasks and outcome measures (Bong 2002; Pajares 1997, 2000; Zimmerman and Bandura 1994). Thus, we hypothesize that guidance views mitigate the impact of low self-efficacy on learning outcomes. Finding ways to diminish the far-reaching effects of low self-efficacy is crucial for science education. Currently, too few students engage with science beyond middle school. Nearly one-third of all students take only a single year of high school science, all but closing the door for future studies in science (Grigg et al. 2006). For many, this choice starts as early as early adolescence and is due to low self-efficacy in science (Leslie et al. 1998).

To investigate whether a MUVE-based guidance system could alleviate some of the differences due to self-efficacy, we conducted an exploratory analysis into the combined impact of guidance use and self-efficacy in scientific inquiry on learning in an educational

MUVE designed to teach scientific inquiry skills. In our analysis, we focus on the following questions:

- How does self-efficacy in scientific inquiry impact use of the guidance system? Do students with lower self-efficacy in scientific inquiry view fewer guidance messages within the presentation of a MUVE-based science curriculum than students with higher self-efficacy in scientific inquiry?
- How does the guidance system mitigate the impact of low self-efficacy in scientific inquiry on learning, if at all? Do students with low self-efficacy in scientific inquiry approach the performance on content tests of students who report high self-efficacy in scientific inquiry as levels of guidance messages increase?

Theoretical underpinnings

Self-efficacy

Self-efficacy refers to a person's belief in his/her ability to perform specific tasks or processes and to achieve designated results (Pajares 1996). As it relates to scientific inquiry, self-efficacy can be defined as a student's perceived capability to perform the tasks involved in scientific inquiry (Ketelhut 2007). It is not a measure of how well students believe they will succeed or how interested they are in the tasks of inquiry; rather, it is a measure of their confidence to conduct the activities at the heart of scientific inquiry.

Researchers have investigated the topic of self-efficacy from various perspectives, describing both the origin of a person's self-efficacy and its effect on behavior. In his seminal work on self-efficacy, Bandura (1977) asserts that a person's belief in his or her own abilities has a powerful impact on whether or not that person initiates "instrumental" actions, and on the extent and length of effort that a person will expend in pursuing those actions. Similarly, Pajares (1996, 2000) contends that self-efficacy affects behavior by regulating the extent of a person's expended effort, their ability to persevere in difficult situations, and their engagement with the task. In those situations, Pajares states that increasing levels of self-efficacy are associated with increasing effort, perseverance, and engagement.

A number of studies have found that high levels of perceived self-efficacy predict performance. In a large meta-analysis of 114 studies into the link between self-efficacy and task performance, Stajkovic and Luthans (1998) found that high levels of self-efficacy were positively correlated with performance of tasks examined in the studies. Lent et al. (1984, 1986) and Lent and Hackett (1987) found that a person's level of self-efficacy impacts both the choice of careers that person chooses to pursue and the level of academic achievement in tasks related to those careers.

Of particular interest to the topic of this study is whether there is any indication of a relationship between level of student self-efficacy and the initial take up and continued use of guidance. The literature in this area is diverse and falls into two main arenas: use of self-regulated learning strategies and help-seeking behavior. Bandura (1986) asserts that self-efficacy positively guides students' choice of self-regulated learning strategies. A good review of the research supporting this tenet can be found in Pajares (2002). The relationship between self-efficacy and help-seeking behavior is equivocal with studies showing widely divergent results regarding whether high or low self-efficacy students are more likely to show help-seeking behaviors and whether this behavior is beneficial for learning or not (for a review see Pajares et al. 2004).

Use of a guidance system embedded in an educational MUVE relies on students' willingness to make use of this novel meta-cognitive strategy, but once they do, the guidance system developed for the current study offers reflective hints rather than directive help. As such, it would appear likely that use of the embedded guidance in a MUVE would more likely reflect the research on self-regulated learning and thus indicate that students with high self-efficacy would be more likely to access the guidance messages than students with low self-efficacy in scientific inquiry. However, since use of the guidance system in the River City MUVE is not directly taught as a learning strategy, it is possible that its appearance in a technology-supported learning environment may mimic the help function often found in traditional educational software packages. If so, then the inconclusive literature on help-seeking behaviors and self-efficacy would be more on point. This study aims to shed light on this dichotomy.

MUVE research

Researchers investigating educational MUVEs as learning environments have explored the design, functionality, and potential impact of such environments on student learning and motivation (e.g. Bers 1999; Bers and Cassell 1998; Corbit and DeVarco 2000; Slator et al. 2004). For example, the research team behind the Quest Atlantis MUVE has published a number of studies about the environment describing its benefits on student engagement and learning (e.g. Barab et al. 2005a,b). Quest Atlantis is a MUVE-based curriculum in which elementary and middle school students can take part in a large number of quests to help the people of Atlantis avoid environmental, moral, and social decay (Barab et al. 2005a, b). In one Quest Atlantis study, Barab et al. (2007) describe a multi-level investigation of the learning benefits associated with a curriculum designed to support scientific inquiry practices situated in realistic, socially relevant issues. Students in the study completed a two-week curriculum designed to support their development of environmental awareness and real-world science inquiry skills while investigating an interactive narrative in the Quest Atlantis environment. Results of the study show promising findings related to student engagement in the MUVE-based curriculum, sophisticated explanations of curricular processes and outcomes, and statistically significant improvement on classroom and standardized assessments.

While research into the Quest Atlantis MUVE has shown promise, engagement among students has been uneven. Lim et al. (2006) studied the levels of engagement exhibited by students participating in virtual scientific inquiry activities in Quest Atlantis. Among the eight participants (11–12 year olds) at a primary school in Singapore, the authors found a low level of engagement as measured by a seven-level "engagement taxonomy." Through interviews with the students and observations of the implementation, the researchers suggested that the biggest contributors to the low engagement were the open-endedness and interactivity of the MUVE, which appeared to distract students from the processes and tasks associated with inquiry.

In her studies of the MOOSE Crossing multi-player environment, Bruckman found similar issues with engagement on the part of children who used the environment to create and share virtual artifacts to learning computer programming (Bruckman 1996, 2000). In one study, Bruckman (2000) performed a portfolio-style assessment of 50 children using MOOSE Crossing to study programming. She found that 40% of the sample group failed to engage in the main curricular task of writing a programming script. Bruckman contended that unevenness in participation and student learning is an inherent by-product of open-ended MUVE-based learning based on a socio-constructivist theoretical framework.

Socio-constructivism and MUVEs

The curricula used with educational MUVEs generally find students learning through open-ended exploration of the software environment with limited embedded guidance support. This kind of unguided collaborative exploration fits well with what Perkins (1991) called "Without Information Given" (WIG) constructivist theory. In this view, students build up a personal understanding of a given topic through self-directed interactions with the content and processes associated with the topic. The curriculum embedded in an early iteration of the River City MUVE reflected this kind of highly unstructured exploratory learning (Dede et al. 2002).

Most recent MUVE-based research focuses on questions of student engagement, socio-cultural interactions, and learning. This research centers on curricula featuring cooperative, open-ended inquiry activities within MUVEs (Barab et al. 2007; Nelson et al. 2005, 2007), and based on a socio-constructivist approach to learning that values free exploration and knowledge building.

There is little agreement among researchers that unguided exploration is the best route to learning. For example, Kirschner et al. (2006) argue that constructivist learning environments with minimal guidance cannot work because they ignore the cognitive processing research into the structure of human memory systems. According to Mayer (2004), in a series of empirical studies dating back to the 1950s, unguided learning has repeatedly been shown to be inferior to guided learning. On the other hand, Flum and Kaplan (2006) describe exploration as a fundamental human trait with developmental, learning, identity formation, and even career benefits. Consequently, they believe the learning should have an "exploratory orientation." However, they acknowledge that the experience of unguided exploration requires tolerance on the part of the individual for ambiguity and uncertainty.

As an alternative to wholly unguided exploration, some researchers suggest that open-ended learning environments such as MUVEs should provide students with tools to build and test hypotheses as scaffolds to the exploration process (Jonassen 1991; Lebow 1993). This type of guidance is called "self-directed," or *reflective*, guidance (Jonassen et al. 1999). Reflective guidance in constructivist environments differs from direct instruction in that it does not provide overt answers or make judgments about particular student actions. Reflective guidance messages instead focus on assisting with student meta-cognition, asking students to reflect upon their learning, describe how they will proceed, and use graphics and/or text to map out their growing understanding (Baylor 2000; Jonassen et al. 1993).

This view of reflective guidance echoes those applied to cognitive "scaffolding," which Puntambekar and Hubscher (2005) define as a system of prompts and hints that support learning. These guidance tools serve to externalize the invisible cognitive processes taking place within each student (Hannafin et al. 1997; Jonassen 1991), and scaffold students' thinking as they develop their own understanding of content present in a given learning environment.

Guidance in MUVEs

In our previous work, one of us has investigated whether the use of embedded reflective guidance in educational MUVEs might offer additional learning benefits over unguided exploration in the virtual worlds. We built a reflective guidance system into the River City educational MUVE, and tested whether use of the guidance led to more effective learning for students. In a study with 272 middle school students completing a MUVE-based science inquiry curriculum, it was found that, while simple exposure to an individualized guidance system in the MUVE had no measurable impact on learning, increased viewing of guidance messages was associated with significantly higher ($p<0.05$) scores from pre- to post-tests on scientific inquiry skills and disease transmission knowledge gained through a MUVE-based curriculum (Nelson 2007).

Although guidance use appeared to help students perform well in the MUVE, there was a great deal of variability in levels of guidance use, and a relatively large (25%) percentage of the students with access to the guidance who never used it. In an effort to account for some of this variation, a number of student demographic variables (gender, SES, prior grades in science, age) and affective measures (prior computer use, experience with gaming, etc.) were examined, but only one was identified as having a statistically significant interaction with guidance use and learning: gender. Although boys and girls both benefited from guidance use, boys viewed significantly fewer hints than girls, and showed lower average gains in content measures across all levels of guidance use.

The River City MUVE project

Our current study is centered on River City, an educational MUVE designed to teach scientific inquiry skills to middle school students. The River City curriculum focuses on skills of hypothesis formation and experimental design, as well as on content related to national standards and assessments in biology and ecology. The main learning goal for students exploring River City is to discover why residents of the virtual town are getting sick (Nelson et al. 2005).

The River City virtual world is set in the late 1800s, and named for the river that runs through most of the town. River City includes explorable digital institutions and buildings such as homes, shops, a library, elementary school, hospital, university, and city hall (see Fig. 1).

Upon entering the city, the students' avatars can interact with computer-based agents (residents of the city), digital objects (pictures and video clips), and the avatars of other students. In exploring, students also encounter visual stimuli such as muddy dirt streets, and auditory stimuli such as the sounds of coughing town residents, which provide tacit clues as to possible causes of illness. Content in an embedded Web browser shifts based on what the student encounters or activates in the virtual environment, such as a dialogue with an agent or historic photos and accompanying text that provide additional information about the town and its residents (see Fig. 2).

Students work in small teams to develop and test hypotheses about why town residents are ill. Water-, air-, and insect-borne diseases are integrated within a curricular framework incorporating historical, social, and geographical content, allowing students to develop and practice the scientific inquiry skills involved in disentangling multi-causal problems embedded within a complex environment (Clarke et al. 2006; Ketelhut et al. 2005). Over the course of the curriculum, students experience a year of virtual time in River City. On each visit, student teams continue to explore River City, working to form a hypothesis.

River City supports hypothesis testing for students by allowing them to change a single factor in one of two identical worlds in order to view the impact, if any, that change had on a particular disease. After student teams design their own hypothesis about the cause of the

Fig. 1 River City

illnesses, they test it by choosing an independent variable to alter. For example, students may decide that cramped living quarters in the public house in River City is a source of illness, and decide to build new housing to lower population density in the "projects." Students then compare the "control" and "experimental" worlds to see if their actions have had any effect on various factors, such as residents' illnesses, water pollution, or number of mosquitoes.

Design and procedure

Research questions

The research questions on which our current analysis is centered are:

1. How does self-efficacy in scientific inquiry impact use of the guidance system? Do students with lower self-efficacy in scientific inquiry view fewer guidance messages within the presentation of a MUVE-based science curriculum than students with higher self-efficacy in scientific inquiry?
2. How does the guidance system mitigate the impact of low self-efficacy in scientific inquiry on learning, if at all? Do students with low self-efficacy in scientific inquiry approach the performance on content tests of students who report high self-efficacy in scientific inquiry as levels of guidance messages increase?

Fig. 2 River City interface

Sample

This paper presents the results of a Fall 2004 implementation in a Mid-Atlantic state with a total sample of 300 middle school students. In the current study, we focus on a sub-population of 96 students (50 boys and 46 girls) who were provided access to an embedded guidance system used in conjunction with the MUVE. This student sample was somewhat homogenous: 6% were eligible for free or reduced lunch and 11% spoke English as a second language. These students were randomly assigned, in teams of three, to the guidance treatment within their seventh-grade science classes, and took part in the River City curriculum with their peers who were randomly assigned to other versions of the curriculum not covered in this study.

Procedures

Students in the treatment group at the center of this study were provided access to an individualized guidance system (IGS) featuring continuously updated guidance links and associated messages. Students worked through the River City curriculum in teams of three, sharing information both in the MUVE and via face-to-face strategy sessions in the classroom. Although students worked in teams, each individual explored River City on his/her own computer. Consequently, each student in the guidance treatment had access to the guidance system on an individual basis.

The guidance in this study was designed as a system of reflective guidance prompts. The system utilized data collected on each student's individual activities to offer reflective questions about the students' data gathering in the MUVE, with the content of the questions based on in-world interactions, such as clicking on pictures, reading signs and charts, and asking questions of computer-based "town residents." To create the guidance system, all objects with which students could interact were tagged with identification codes. When a student clicked on an object or "spoke" to a *River City* citizen, a record of the event was stored in a database. A guidance model was triggered after each student interaction in the MUVE. A subset of these interactions was associated with guidance scripts. These scripts then prompted the appearance of "hint buttons" that, when clicked, displayed reflective questions to students.

Students with access to this guidance system could view three hints per pre-defined information hot-spot object in River City. Whenever students clicked on a specially tagged object inside the virtual world, the guidance system would flash alternating colors to indicate that new hints were available. The IGS did not automatically show specific guidance messages but instead displayed hint buttons linked to messages (see Fig. 3). To view guidance messages, students needed to click on these hint buttons. In this way, we were able to track if and when students chose to access the guidance messages and which messages they saw.

Students spent approximately the next 12 days participating in the project as part of their science classes. The first four of these days were spent in the computer lab exploring and gathering information from four sequential versions, worlds, of River City itself. During visits 1 and 2, students were encouraged to explore River City; to interact with the computerized residents, the library books, the admissions record and the Smithsonian artifacts. During visits 3 and 4, students continued to explore and were also given access to the water and bug sampling stations. These four days were followed by two days of face-to-face experimental design group work in the regular classroom. Students then returned to the computer lab and re-entered River City to gather information to test their hypothesis, first from the control world and then from the experimental world, which differed by one factor that each team of students chose in advance. Student teams then spent two days in the classroom analyzing their data. After this, students were asked to write a report to the Mayor of the town in which they discussed their hypothesis, experimental design, results and recommendations for solving the city's health problem. Finally, a mini-research conference was held in each classroom to allow student teams to report on their findings.

Measures

Both qualitative and quantitative data were collected from students over the implementation period. Pre- and post-intervention, the students completed an affective measure that consisted of various subscales from three different surveys, Self-Efficacy in Technology and Science (SETS; Ketelhut 2005), Patterns for Adaptive Learning Survey (Midgley et al. 2000), and the Test of Science Related Attitudes (Fraser 1981). Of particular interest to this study was the subscale used from the SETS: self-efficacy in scientific inquiry. This subscale contains 12 items (see Appendix) with each rated on a scale from 1 (low) to 5 (high). Overall scores are computed by averaging the student's responses across the 12 items of the subscale, with high scores representing high self-efficacy. The measure has an estimated internal consistency reliability of 0.86 in a population of middle school students.

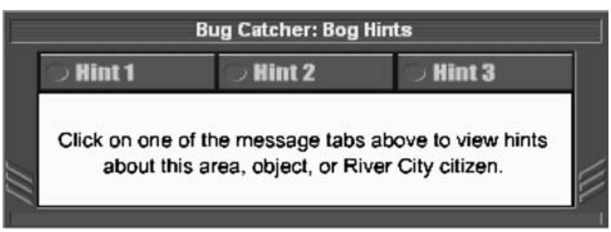


Fig. 3 Guidance system

To assess understanding and content knowledge (science inquiry skills, science process skills, biology), we administered a 30-question content test, pre- and post-intervention with an internal consistency reliability of 0.80 in a middle school population. These tests were designed to evaluate whether and to what extent participating students had increased their knowledge of the desired outcomes of the curriculum, that is, an understanding of and ability to apply scientific inquiry skills to investigate a real-world problem, and an understanding of various methods of disease transmission.

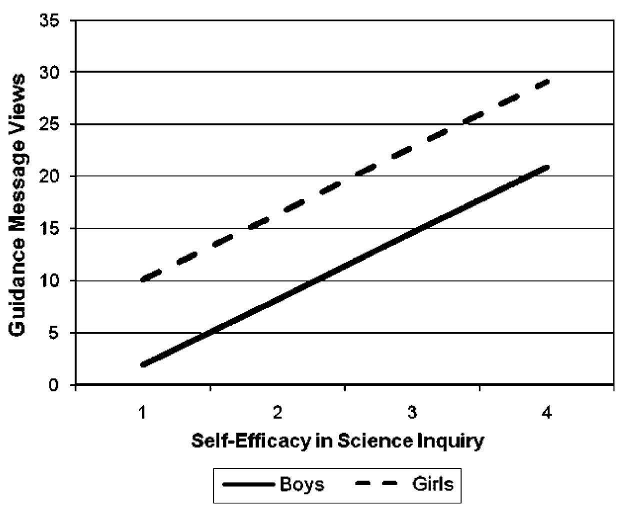
Findings

The quantitative data were analyzed with SAS, using a significance level of $p \leq 0.05$; checks for linearity, normality and homoscedasticity were performed at intervals with no violations found. Our first research question asks whether students with low self-efficacy in scientific inquiry made less use of the embedded guidance system than those with higher self-efficacy. The answer is yes. In this study, we found that students with low self-efficacy in scientific inquiry viewed significantly fewer guidance messages ($p<0.05$) than their peers with higher initial self-efficacy. In addition, it was found that boys viewed significantly fewer messages than girls ($p<0.05$), overall. This gender difference was still evident when initial self-efficacy level was taken into account. In other words, boys across a range of initial science self-efficacy scores viewed fewer guidance messages than girls at the same level ($p<0.05$). This relationship can be seen in Fig. 4.

To assess research question 2, we first regressed student posttest scores on levels of use of the individualized guidance system and pretest scores. In this analysis, we found that viewing guidance messages had a significant positive impact on posttest scores ($p<0.01$). In other words, holding pretest scores constant, students who viewed more guidance messages outperformed students who viewed fewer.

To investigate whether the guidance system could mitigate the effect of low self-efficacy students on learning, we added initial level of self-efficacy in scientific inquiry to the

Fig. 4 Guidance message views predicted by self-efficacy in scientific inquiry, by gender (n=94)



model. Since one of us (Ketelhut 2007) has shown that self-efficacy and pre-test science content scores have overlapping effects in educational MUVE implementations, pre-test scores were not included in this model. Here we found that students with lower initial self-efficacy scores did not perform as well as their higher efficacy peers at each level of guidance message viewing. In other words, students with lower initial self-efficacy scores performed less well than higher self-efficacy students across a spectrum of guidance message viewing. Although use of the guidance system helped low self-efficacy students perform better, guidance use was not able to bridge the self-efficacy gap in learning outcomes. Figure 5 shows this relationship. However, it is important to note that guidance viewing did mitigate the impact of low self-efficacy on learning to a small degree. Students with low self-efficacy who viewed a large number of guidance messages did as well as students with high self-efficacy who viewed no messages. Thus, while it cannot level the playing field evenly, the guidance system did improve outcomes.

Discussion and conclusion

From our prior research we learned that use of embedded guidance in the River City MUVE could help improve student science content scores. However, we also found that viewing levels of guidance messages varied widely among students with access to them. In this study, we sought to account for some of that variability by examining the role that student self-efficacy in scientific inquiry might play on guidance use. Confirming what the literature says regarding behavioral choices and self-efficacy, we can see from the results that initial self-efficacy on entry into the River City MUVE affects overall how students make use of the embedded guidance system, with low self-efficacy students viewing fewer guidance messages on average than their higher self-efficacy peers. Since it has been shown that use of the guidance system in River City can improve academic outcomes (Nelson

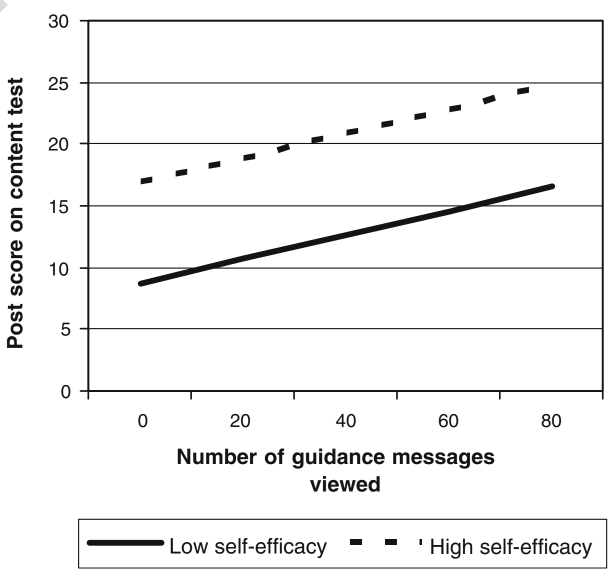


Fig. 5 The effect of self-efficacy and guidance views on post test outcomes (n=94)

2007), the lower levels of guidance use among students with low self-efficacy may be handicapping their learning in the environment. River City has been designed since its inception with a goal of providing a motivating, engaging environment for under-achieving students. Unless a way to boost guidance use by low self-efficacy students can be found, implementations of a curriculum with a strong pedagogical bent toward helping these students close the learning gap with their higher-achieving classmates will not reach its mark.

However, what is not clear from this analysis is whether the relationship between seeking guidance and self-efficacy is constant throughout the project or varies over time and with exposure. Is it possible that continued exposure to guidance hints helps change behavior so that there is little difference in behavior between low and high self-efficacy students? If so, could that mean that given time all students might begin to take advantage of the learning process?

There is research that would indicate that this is the case. Schunk (1983, 1987) has found that self-efficacy appears to rise when students are provided with just-in-time feedback, help, and signals that success was due to their own hard work. These are all features that can be found in an embedded guidance system. In addition, one of us conducted an in-depth investigation with this same student population examined in the current study and curriculum implementation to investigate the influence of student self-efficacy in scientific inquiry on data-gathering behavior in the MUVE over time (Ketelhut 2007). In that study, student self-efficacy predicted for behavior initially, but by the fourth visit to the River City virtual town, the effect of self-efficacy on interactions with the environment had disappeared. In other words, students with low self-efficacy prior to the start of the study behaved no differently than students with high self-efficacy by their fourth visit to the environment.

Our other major finding is that guidance viewing can only mitigate the impact of low self-efficacy on learning in a very moderate way. However, this is an important finding for the community of researchers on help-seeking. Recall that the literature on whether help-seeking behaviors improve learning is equivocal. Yet, in our study, we discovered that not only did guidance views improve outcomes, it did so for students at all levels of self-efficacy.

Of interest is our un-hypothesized finding regarding gender—that girls across all levels of self-efficacy accessed more guidance messages than boys. In its report on educating girls in the computer age, The American Association of University Women (AAUW) educational foundation lists a number of design suggestions for computer games that will hold appeal for girls, including rich narrative, customizable avatars, opportunities for collaboration and communication, and opportunities for positive social action (American Association of University Women 2000).

Because MUVEs support many of the features suggested by the AAUW report as useful for "girl-friendly" software, it is perhaps not surprising that our findings indicated relatively higher levels of guidance use among girls. Past MUVE studies, including our own described earlier (Nelson 2007), have indicated that MUVEs are supportive of high levels of participation, motivation, and learning outcomes among girls. One such study on gender and programming achievement in a text-based MUVE found that girls spent significantly more time than boys communicating with others in the environment (Bruckman et al. 2002). Barab and his colleagues conducted an extensive analysis of gender participation in the Quest Atlantis (QA) MUVE (Socially-Responsive Design Group 2004) and found no differences in terms of overall participation rates in the MUVE between boys and girls. However, looking specifically at participation as reflected by online communication, it was found that girls used chat more than boys ($p<0.01$) and sent more e-mail messages ($p<0.01$) than boys. In addition, Barab's group found that girls wrote more in their online notebooks when completing quests and engaged in longer metacognitive reflections about their work in the MUVE. This finding, although not a focus of the current study, is one worth pursuing in future MUVE research.

Our overall findings indicate that future research should investigate methods through which use of guidance embedded in educational MUVEs can be better facilitated for low self-efficacy students. We suggest that future studies draw from literature on guidance- and help-seeking strategies to find ways to better support guidance uptake and continued use by low self-efficacy students. Another tack we suggest for future research is to investigate methods that can be used by MUVE curriculum designers to help improve the self-efficacy of learners as a result of their interactions in these kinds of learning environments. Raising self-efficacy of all learners is vital, as the effects of self-efficacy are far ranging. For example, students with low self-efficacy in a given area are less likely to choose a career in that domain (Lopez and Lent 1992). With well-designed educational MUVE-based curricula incorporating individualized guidance and engaging inquiry, we hope that all learners can better understand and apply principles of real-world science inquiry.

Appendix

Self-efficacy in Scientific Inquiry

I can write an introduction to a lab report.

I can use graphs to show what I found out in my experiment.

It is hard for me to write a report about an experiment.

I know how to use the scientific method to solve problems.

It is hard for me to look at the results of an experiment and tell what they mean.

When I do an experiment, it is hard for me to figure out how the data I collected answers the question.

When I do my work in science class, I am able to find the important ideas.

Once I have a question, it is hard for me to design an experiment to test it.

I can design an experiment to test my ideas.

I have trouble figuring out the main ideas of what my science teacher is teaching us.

I can tell the difference between observations and conclusions in a story.

It is easy for me to make a graph of my data.

References

American Association of University Women (2000). *Tech-savvy: Educating girls in the new computer age.* Washington, DC: AAUW Educational Foundation.

Bandura, A. (1977). Self-efficacy: towards a unifying theory of behavioural change. *Psychological Review, 84*, 191–215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Barab, S., Arici, A., & Jackson, C. (2005a). Eat your vegetables and do your homework: a design based investigation of enjoyment and meaning in learning. *Educational Technology, 45*(1), 15–20.

Barab, S., Sadler, T., Heiselt, C., Hickey, D., & Zuiker, S. (2007). Relating narrative, inquiry, and inscriptions: supporting consequential play. *Journal of Science Education and Technology, 16*(1), 59–82.

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzan, H. (2005b). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research & Development*, *53*(1), 86–107.
Baylor, A. (2000). Beyond butlers: intelligent agents as mentors. *Journal of Educational Computing Research*, *22*(4), 373–382.
Bers, M. U. (1999). Zora: A graphical multi-user environment to share stories about the self. In C. Hoadley & J. Roschelle (Eds.), *Computer support for collaborative learning: Designing new media for a new millennium. Proceedings of CSCL 1999*. Palo Alto, CA, USA.
Bers, M. U., & Cassell, J. (1998). Interactive storytelling systems for children: using technology to explore language and identity. *Journal of Interactive Learning Research*, *9*(2), 183–215.
Bong, M. (2002). Measuring self-efficacy: multitrait–multimethod comparison of scaling procedures. *Applied Measurement in Education*, *15*(2), 143–171.
Bruckman, A. (1996). Finding one's own space in cyberspace. *Technology Review*, *99*(1), 48–54.
Bruckman, A. (2000). *Uneven achievement in a constructivist learning environment*. Paper presented at the International Conference on Learning Sciences, Ann Arbor, MI.
Bruckman, A., Jensen, C., & DeBonte, A. (2002). Gender and programming achievement in a CSCL environment. In G. Stahl (Eds.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of CSCL 2002*. Boulder, CO, USA. Mahwah, NJ: Erlbaum.
Clarke, J., & Dede, C. (2005). *Making learning meaningful: An exploratory study of using multi-user environments (MUVEs) in middle school science*. Paper presented at the American Educational Research Association Conference, Montreal, Canada.
Clarke, J., Dede, C., Ketelhut, D. J., & Nelson, B. (2006). A design-based research strategy to promote scalability for educational innovations. *Educational Technology*, *46*(3), 27–36.
Corbit, M., & DeVarco, B. (2000). *SciCentr and BioLearn: Two 3D implementations of CVE science museums*. Paper presented at the Third International Conference on Collaborative Virtual Environments, San Francisco.
Dede, C., Ketelhut, D. J., & Ruess, K. (2002). Motivation, usability, and learning outcomes in a prototype museum-based multi-user virtual environment. In P. Bell, R. Stevens & T. Satwicz (Eds.), *Keeping learning complex: The proceedings of the Fifth International Conference of the Learning Sciences*. Mahwah, NJ: Erlbaum.
Flum, H., & Kaplan, A. (2006). Exploratory orientation as an educational goal. *Educational Psychologist*, *41*(2), 99–110.
Fraser, B. (1981). *TOSRA: Test of science related attitudes*. Hawthorne, VIC: Australian Council for Educational Research.
Galarneau, L & Zibit, M (2008). Online games for 21st century skills in games and simulations in online learning: Research and development frameworks, eds Gibson, Aldrich, Prensky, in press Q1
Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York, NY: Palgrave.
Grigg, W., Lauko, M., & Brockway, D. (2006). *The nation's report card: Science 2005 (NCES 2006–466). U.S. Department of Education, National Center for Education Statistics*. Washington, D.C.: U.S. Government Printing Office.
Hannafin, M. J., Hannafin, K. M., Land, S. M., & Oliver, K. (1997). Grounded practice and the design of constructivist learning environments. *Educational Technology Research & Development*, *45*(3), 101–117.
Jonassen, D. H. (1991). Objectivism versus constructivism: do we need a new philosophical paradigm? *Educational Technology Research & Development*, *39*(3), 5–14.
Jonassen, D. H., Peck, K. L., & Wilson, B. G. (1999). *Learning with technology: A constructivist perspective*. Upper Saddle River, NJ: Merrill, Prentice Hall.
Jonassen, D. H., Wilson, B. G., Wang, S., & Grabinger, R. S. (1993). Constructivist uses of expert systems to support learning. *Journal of Computer-Based Instruction*, *20*(3), 86–94.
Ketelhut, D. J. (2005). *Assessing science self-efficacy in a virtual environment: A Measurement Pilot*. Paper presented at the National Association of Research in Science Teaching Conference, Dallas, April.
Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: an exploratory investigation in River City, a multi-user virtual environment. *The Journal of Science Education and Technology*, *16*(1), 99–111.
Ketelhut, D. J., Clarke, J., Dede, C., Nelson, B., & Bowman, C. (2005). Extending library services through emerging interactive media. *Knowledge Quest*, *34*(1), 29–32.
Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.
Lebow, D. (1993). Constructivist values for instructional systems design: five principles toward a new mindset. *Educational Technology Research and Development*, *41*(3), 4–16.
Lent, R. W., Brown, S. D., & Larkin, K. C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology*, *31*, 356–362.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology*, *33*, 265–269.
Lent, R. W., & Hackett, G. (1987). Career self-efficacy: empirical status and future directions. *Journal of Vocational Behavior*, *30*, 347–382.
Leslie, L. L., McClure, G. T., & Oaxaca, R. L. (1998). Women and minorities in science and engineering: a life sequence analysis. *The Journal of Higher Education*, *69*(3), 239–276.
Lim, C. P., Nonis, D., & Hedberg, J. (2006). Gaming in a 3D multiuser virtual environment: engaging students in science lessons. *British Journal of Educational Technology*, *37*(2), 211–231.
Lopez, F. G., & Lent, R. W. (1992). Sources of mathematics self-efficacy in high school students. *The Career Development Quarterly*, *41*, 3–12.
Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *Educational Psychologist*, *59*, 14–19.
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: Observe, interact, change, learn*. Washington, D.C.: National Academy Press.
Nelson, B. (2007). Exploring the use of individualized, reflective guidance in an educational multi-user virtual environment. *The Journal of Science Education and Technology*, *16*(1), 83–97.
Nelson, B., Ketelhut, D. J., Clarke, J., Bowman, C., & Dede, C. (2005). Design-based research strategies for developing a scientific inquiry curriculum in a multi-user virtual environment. *Educational Technology*, *45*(1), 21–34.
Nelson, B., Ketelhut, D. J., Clarke, J., Dieterle, E., Dede, C., & Erlandson, B. (2007). Robust design strategies for scaling educational innovations: The River City case study. In B. E. Shelton & D. A. Wiley (Eds.), *The design and use of simulation computer games in education* (pp. 219–242). Rotterdam: Sense.
Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, *66*, 543–578.
Pajares, F. (1997). Current directions in self-effiacacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10 (pp. 1–49). Greenwich, Ct: JAI.
Pajares, F. (2000). *Schooling in America: Myths, mixed messages, and good intentions (Lecture)*. Cannon Chapel: Emory University.
Pajares, F. (2002). *Overview of social cognitive theory and of self-efficacy*. Retrieved September 5, 2007, from http://www.emory.edu/EDUCATION/mfp/eff.html.
Pajares, F., Cheong, Y. F., & Oberman, P. (2004). Psychometric analyses of computer science help-seeking scales. *Educational and Psychological Measurement*, *64*, 496–513.
Perkins, D. N. (1991). Technology meets constructivism: do they make a marriage? *Educational Technology*, *31*(5), 18–23.
Pintrich, P. R., & Groot, E. V. D. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33–40.
Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: what have we gained and what have we missed? *Educational Psychologist*, *40*(1), 1–12.
Schunk, D. H. (1983). Ability versus effort attributional feedback: differential effects on self-efficacy and achievement. *Journal of Educational Psychology*, *75*, 848–856.
Schunk, D. H. (1987). Peer models and children's behavioural change. *Review of Educational Research*, *57*, 149–174.
Shaffer, D. W. (2006). *How computer games help children learn*. New York: Palgrave.
Slator, B. M., Hill, C., & Del Val, D. (2004). Teaching computer science with virtual worlds. *IEEE Transactions on Education*, *47*(2), 269–275.
Socially-Responsive Design Group (2004). *Creating a socially-responsive play space for learning: Something for girls and boys*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: a meta-analysis. *Psychological Bulletin*, *124*, 240–261.
Steinkuehler, C. A. (2004). Learning in massively multiplayer online games. In Y. B. Kafai, W.A. Sandoval, N. Enyedy, A. S. Nixon & F. Herrera (Eds.), *Proceedings of the Sixth International Conference of the Learning Sciences* (pp. 521–528). Mahwah, NJ: Erlbaum.
Steinkuehler, C., & Chmiel, M. (2006). Fostering scientific habits of mind in the context of online play. In S. A. Barab, K. E. Hay, N. B. Songer & D. T. Hickey (Eds.), *Proceedings of the International Conference of the Learning Sciences* (pp. 723–729). Mahwah NJ: Erlbaum.
Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal*, *31*, 845–862.