The Effectiveness of Learning Geography Using Computer-based Games

Jerry Prell, William R. Nelson, and John Foshay,
Central Connecticut State University

Eighty, seventh grade students attending a suburban middle school in southern Connecticut participated in the study evaluating the effectiveness of computer based geography games on student motivation and achievement. Using Connecticut Mastery Test (CMT) Scores and baseline U.S.A scores as criteria to develop matched pairs, students were divided so that the control group played Sheppardsoftware.com's States Level 3 and the experimental group played States Level 6. According to the data collected, States Level 3 and States Level 6 are significantly effective instructional tools that students enjoyed playing. CMT scores in spatial reasoning did not prove strong predictors of students' performance whereas U.S.A. baseline testing was a strong indicator in two of three categories. Results support the use of computer-based games to increase student motivation and social studies achievement.
        Keywords: Geography, Computer-based instruction, US Maps

The No Child Left Behind Act (NCLB) is focused on high stakes testing in both language arts and math and has led to increasing pressure on districts to meet annual yearly progress goals, thus, profoundly affecting direct social studies instruction at both the elementary and secondary level. The overall decrease in time dedicated to social studies has already begun to be manifested in weaknesses in students' social studies knowledge base. Many incoming middle school students cannot accurately produce or identify the seven continents and major oceans as a part of their skill set. This lack of knowledge is a troubling sign for our upcoming generation of students as they actively prepare to become members of the ever-increasing globally connected economy. As the social studies field faces increased marginalization by administration due largely in part to NCLB pressures, teachers must use this

limited amount of instructional time to the maximum extent possible.

Additionally, with the reauthorization of the Individuals with Disabilities Education Act (IDEA) in 2004 the federal government established that a Response To Intervention (RTI) approach to the identification of students with learning disabilities could be utilized. RTI encourages the effective differentiation of instruction in order to implement its three-tiered structure. Differentiation for all levels of learners is key to the success of this tiered intervention model, especially in tier one. Differentiated lessons for all learners can reduce educational gaps resulting from poor, mismatched or inadequate instruction. When students do not show improvement in tier one, tiers two and three can be applied using increasingly frequent interventions implemented to smaller groups. Clearly, information regarding the effectiveness of computer games when applied to geography curriculums matches well with goals outlined by RTI. Gaining insights into what works best for some students as opposed to others, will support the RTI process and provide for future applications in other classrooms.

General education social studies teachers at the middle school level when attempting to implement World Geography Curriculums are often confronted with diverse map and literacy skills among their general students and students with disabilities found in their classrooms. It is not uncommon for large gaps to exist between students who have varying degrees of strengths and weaknesses in their abilities to read, use and produce maps. Differentiating instruction related to map literacy skills is quite challenging but is essential. Student frustration escalates when lessons are not appropriately differentiated to accommodate for varied skill levels. Clearly, when learning objectives are either too challenging or too simple student learning motivation is sacrificed.

It is difficult to design and implement effectively differentiated lessons for middle school students with such a wide distribution of map literacy skills. For example, challenges arise when teaching students to learn the names of the U.S.A.'s 50 states, recognize their shapes, and correctly identify their locations on a map. While some students can easily store and recall the names of each state, a few will struggle to complete the activity without a list. In addition, some students require even more clues; for them, providing the first letter of each state name within its shape works to help limit the possible choices. For some students 50 states might be too large of a group, and so they may need to have them broken down into smaller regions or groups. When differentiating the content, it is also important to take into account that some students will finish the task faster then others.. Keeping students motivated and engaged is challenging, in a classroom where some students are experiencing great success with maps while others are showing little improvement in map skills even with much practice.

The noticeable difference between students with higher and lower aptitudes for map literacy skills likely exists either because there are gaps in students' learning due to poor or inadequate instruction or that some students naturally have better spatial reasoning and visual memory skills than others. Social studies teacher begin each year with little knowledge of their

students' map literacy aptitude and less knowledge of their spatial reasoning and visual memory skills. Since many teachers already apply differentiated instruction, it may be helpful practice to gain a better understanding of students' spatial reasoning and visual memory skills through and assessment before direct instruction in mapping occurs. Thus, by first acquiring a baseline for student skills, and designing instruction using this knowledge to inform instructional efforts.

Furthermore, social studies teacher in planning instruction must consider influencing student motivation for learning the material. The implementation of computer based software games as an intervention to help students acquire geographic knowledge could increase student learning motivation and enthusiasm. The competition and reinforcement of computer based software games could lead to increased guided practice involvement for students.

To effectively teach a geography course without using maps would be impossible. When addressing the importance of maps in social studies, Bednarz, Acheson, and Bednarz (2006) stated, "Maps are not the whole of geography, but there can be no geography without maps" (p. 398). The importance of map competency for contemporary students must not be underestimated. Due to a variety of factors, including greater technological capabilities as well as lesser printing and production costs, the overall quality and quantity of maps have continued to increase since 1990 (Bednarz, et al., 2006). Without guided practice necessary to develop and refine the spatial skills necessary to utilize maps as resources, especially beginning in early

grades, students may struggle to find success with geographic related concepts in our ever changing, globally connected world (Gersmehl & Gersmehl, 2007).

The Committee on the Support for Thinking Spatially of the National Research Council (CSTS, 2005) defined the process of spatial reasoning as having three dimensions – concepts of space, tools of representation and processes of reasoning. Concepts of space provide the framework in which representations can be made, and thus, from that point stored, analyzed, comprehended and shared by individuals (CSTS, 2005). According to Gersmehl and Gersmehl (2007), children are capable of using and developing spatial thinking skills at a very young age. Attributes such as skillful painting, drawing, storytelling, and use of pictures, video, diagrams or maps are signs that an individual has strengths in spatial abilities (Armstrong, 2001). Furthermore, a study by Stavridou and Kakana (2008) has shown that these types of strengths in spatial abilities may correlate with enjoyment and success in both mathematics and science. Gersmehl and Gersmehl (2007) suggested adult intervention is helpful for individuals to further enhance their development of spatial abilities and Bednarz et al. (2006) posited that without formal instruction individuals' spatial abilities will eventually plateau. Hence, the ability to think in spatial terms appears to be an innate ability that can be improved with instruction and practice throughout the course of a lifetime (Gersmehl & Gersmehl, 2007).

The ways in which students learn, practice and use their spatial skills in geography are undergoing a rapid transformation. According to the Rediscovering Geography Committee of

the National Research Council (RGC, 1997), maps are traditionally viewed as two-dimensional pieces of paper, but recent developments in technology have helped to render this viewpoint as obsolete. In an analysis of mapping trends, the RGC (1997) stated, "The modern map is a dynamic and multidimensional product that exists in digital form, opening up new areas of research and application for geographic investigation" (p. 3). Across the geographic field, increased exposure to maps via electronic media has led researchers such as Bednarz et al. (2006) to suggest that a greater focus on technological literacy in the classroom is necessary for both teachers and students.

Technological advances have made learning through media more accessible to all students and, consequently, the educational community has responded by shifting some delivery of instruction from traditional methods such as textbooks to media such as digital games (Annetta, 2008). In fact, Hong, Cheng, Hwang, Lee and Chang (2009) reported, "Recently, digital games have played an increasingly crucial role in developing intelligence in young learners" (p. 431). The popularity of digital games commercially has caused the educational community to examine their specific potential as in instructional tool (Moreno-Ger, Burgos, & Torrente, 2009). Annetta (2008) referred to today's K-20 students as the *net generation*, a group who now views socialization and entertainment in a new way, and argued that 20th century instructional materials should not be used to teach 21st century skills.

The benefit of games as an educational tool has long been studied (Moreno-Ger et al., 2009). Piaget (as cited in Hong et al., 2009) argued a correlation existed between game behavior and learning. In attempts to determine the extent of this relationship, contemporary researchers have recently begun assessing the effectiveness of digital game implementation during academic activities. Digital game use has shown promise by providing students with opportunities to engage in learning and complete assessments in ways that are both engaging and motivating (Salend, 2009). In some cases, students' higher levels of motivation are quite clear. For instance, students who participated in educational digital games during one study showed an increase in grades across all subjects and were more likely to get involved in learning activities related to areas studied while playing the games during hours outside of school (Squire, DeVane, & Durga, 2008).

As the popularity of digital game use in the classroom has grown, much research and analysis has been devoted to the existence of barriers to successful implementation. Challenges to successful curriculum integration, the necessity to meet technical and logistical requirements, and lack of proper teacher training have been raised as potential roadblocks (Kebritchi, Hirumi, Kappers, & Henry, 2008). Also, Liu and Johnson (2005) reported that, oftentimes, available games do not match up well with both students' and educators' needs. Surprisingly, students' overuse of technology may also be a problem. One recent study has found that daily optimal computer use for students is around one hour per day, meaning too much extra exposure in the classroom may not yield additional benefits (Sang, Brescia, & Kissinger, 2009).

There are, however, many

suggestions about how to overcome these obstacles.

In response to criticism regarding educational fit, researchers have been working to develop more effective and appropriate games that are standards based (Liu & Johnson, 2005). Hong et al. (2009) proposed a framework for an educational value index to assess the value of digital games and emphasized the need for educators, researchers and game developers to collaborate in order to produce games which are effective learning tools. When addressing how teachers may practically implement digital games into the classroom, Salend (2009) stresses that educators should proceed gradually, bridge the digital divide, teach students to be good digital citizens, and keep up to date with new technologies and assessment strategies. Thus, it falls upon educators to analyze the relative strengths and weaknesses of available digital games and design learning applications which best match their students' needs in the intended setting (Liu & Johnson, 2005).

Some studies have yielded promising results for educators who have tried to incorporate digital games into social studies classes. When using historical simulation digital games in an after school setting for one year, Squire et al. (2008) found that participants' interest in social studies increased. In addition, Squire et al. (2008) suggested that gaming may serve best as an introductory "hook" because it led students to become interested into "more academically valued practices such as reading or watching documentaries" (p. 249). In another study, analysis of pre and post achievement tests showed 4th and 5th grade geography students made significant learning gains by participating in game-based learning (Tüzün, Yilmaz-Soylu, Karakus, Inal, & Kizilkaya, 2009). Furthermore, Tüzün and colleagues (2009) reported that students who played the digital games were more independent learners who displayed increased levels of intrinsic motivation and decreased focus on getting grades. In both studies, researchers worked to match the capabilities of the games each selected with the needs of their participants. Squire et al. (2008) specifically chose *Civilization III* as an educational game aimed at motivating students to study ancient history whereas a 3- dimensional game was developed specifically for the students to learn about continents and countries in the Tüzün et al (2009) study. Salend (2009) captured the importance of this process when he stated that educators should "carefully evaluate the various technologies to identify those most effective, equitable, and appropriate for use by students and teachers, and determine the extent to which the use of technology-based assessment strategies align with their instructional program and curricular goals" (p. 58).

Little research has been aimed at measuring the possible correlation between strong spatial abilities (aptitude) and performance on geographic mapping tasks (treatment). However, there have been several studies measuring performance on an assessment after using multimedia instructional techniques that caused a split of attention during tasks requiring spatial ability. Coluccia, Bosco and Brandimonte (2007) studied the role of visuo-spatial working memory (VSWM) in map learning. Participants in one part of this study were asked to read a map, and then later, reproduce the map by drawing it. The control group studied for five

minutes with no distractions, a second group studied while tapping a sequence of numbers on a keypad as a spatial secondary task, and a third group which said aloud a sequence of syllables as verbal secondary task. Results showed that map drawing performance was significantly impacted by the spatial distractions but not the verbal distractions (Coluccia et al., 2007). Another study by Seufert, Schütze, and Brünken (2008) tested the modality effect in multimedia, which states a picture labeled with text should have the text read auditorily to avoid a split of attention. In part of the study, one group was taught test material with auditory support while the other was not. Results showed that learners with higher memory strategy skills could better compensate during visual only instruction and scored similarly to those who received auditory support. However, learners with lower skills performed significantly lower under visual only conditions than those who received auditory support (Suefert et al., 2008). Results from both studies showed that secondary spatial tasks can cause distractions, but Suefert et al. (2008) suggested those with stronger spatial abilities can better cope with those distractions and are more likely to be successful.

Other studies have further examined relationships between spatial abilities and performance on spatial assessment tasks. Using data from pretest and posttest scores as evidence, Onyancha, Derov and Kinsey (2009) reported that college students training to work in the engineering field significantly benefited from targeted spatial ability training using CAD software over a period of time compared to those who did not participate in the program. In fact, the experimental group, determined according to their scores of less than 60% on the pretest, benefited so much from the spatial training that the group improved their scores on the posttest to roughly the level of those students who also received training but had scored 80% on the pretest. In another study, Keehner, Hegarty, Cohen, Khooshabeh, and Montello (2008) found that viewing the most important information relevant to completing a spatial task was an important factor, regardless of whether it was obtained interactively or not. Interestingly, the study also reported that spatial ability made an independent contribution to task performance in that students with high spatial ability were able to benefit from external visualizations rather than students with low spatial ability using the external visualizations to compensate, like a prosthetic, in their area of weakness (Keehner et. al, 2008).

Relative to geography, some studies have explored the impact of spatial learning strategies and map performance without fully considering spatial ability level as a factor. For example, Friedman (2009) tested college students on their ability to estimate locations of North American cities. As part of the study, students worked at computers and were required to drag and drop an *X* on the screen as close as they possibly could to the correct location of the city. Varying amounts of memory aids, or cues, were provided to three different groups. Data from the results suggested accurate spatial cues helped participants make better estimates by limiting the amount of response error they could display. Furthermore, Carpenter and Pashler (2007) demonstrated that the testing effect, the process of repetitive testing which has

been shown to improve performance in verbal memory tasks, could be applied to improving spatial ability as well. All participants in this study were college students who had been shown a complete map and assessed on their ability to correctly reproduce it. Those who were taught the map through repetitive testing made significantly larger gains than those who had extra opportunities and extended time to restudy the material (Carpenter & Pashler, 2007). Similarly, Smith, Morey and Tjoe (2007) presented findings that college students who practiced learning a computer game using feature masking, when visual information is either removed or covered up, used visual imagery strategies to a significantly greater degree than those who did not experience feature masking. The use of visual imagery strategies predicted success in both the experimental and control groups. Smith et. al. (2007) described feature masking as a valuable instructional too for educators in fields such as geography because it is a scaffolded process in which students begin by needing more sensory support when they have less skill and can be "spatially weaned" as they develop more skill (p. 360).

Research Questions

The research questions were: a) does the implementation of computer-based geography games significantly improve students' acquisition and motivation of knowledge regarding the 50 states of the USA and b) does a significant correlation exist between students' CMT spatial reasoning skills, baseline USA 50 state scores.and performance on the computer-based geography games.

Methodology
Participants

Participants included seventh grade students attending a suburban middle school in southern Connecticut. The sample was taken from a pool of 104 students from four geography classes. Each class had 26 members. The final group of 80 participants consisted of 45 males and 35 females with a mean age of 13.2 years. Forty sets of pairs were formed using a matched pairs technique. Each pair was then randomly split into a control group and an experimental group. The control group consisted of 24 males and 16 females with a mean age of 13.0 years, while the experimental group included 21 males and 19 females with a mean age of 13.4 years. Students arrived in seventh grade having been similarly exposed to geographical content in sixth grade during their social studies coursework on ancient civilizations such as Egypt, Mesopotamia and Rome. See Table 1 for a presentation of participants with disabilities.

Table 1
Demographic Information of Participants With Disabilities

Disability Label	USA Map Pretest Score	Group
Learning Disability	18	Experimental
Learning Disability	29	Experimental
Learning Disability	3	Experimental
Learning Disability	4	Experimental
Learning Disability	9	Experimental
Learning Disability	7	Experimental
Autism Spectrum	50	Control
Learning Disability	15	Control
Learning Disability	5	Control
Learning Disability	11	Control
Learning Disability	9	Control

Note. USA Map Pretest Score column refers to number of states identified correctly.

Materials

To gather baseline data for this study, an initial examination on the 50 States and selected Connecticut Mastery Test (CMT) scores were used. Then, two computer-based map games were chosen as intervention and assessment tools. Lastly, a writing prompt was selected and used to gauge student ratings of the games.

Standardized, norm-referenced CMT tests are administered yearly to students in Connecticut. Participants' last completed CMT testing during sixth grade approximately one year prior to this investigation. Data from Math Strand 18, Spatial Relationships, was recorded for use in pairing students by ability level. Then, a baseline examination was developed to ascertain students' prior knowledge relating to the 50 States. This fill-in-the-blank examination required students to correctly write the name of each state according to the corresponding number on a given map without a word bank. There was no penalty for misspelling.

Subsequently, two educational computer games produced by Sheppardsoftware.com (2007) were used as intervention and assessment tools. Both games were located under U.S.A. Games. Students worked with States Level 3 while the other was States Level 6. States Level 3 required players to drag each of the 50 states to the correct location and drop it there, like putting together pieces of a puzzle. States Level 6 was predicated on the same requirements, but when placed correctly, states disappear. Thus, the game became progressively more difficult due to a decrease in availability of visual cues. Directions were created for students to read in the computer lab before they began playing the game. At the end of the directions, a hyperlink was embedded as a portal to help students access the correct game. Lastly, in order to gauge students' interest the games, a writing prompt was developed. The prompt began, "*A friend comes up to you at lunch and asks, "Did you play a game in geography class today?*" Then, a question was posed. "In your response,

will you tell them that the game was really fun, just all right or not very fun at all?" Students were asked to include at least three reasons to support their answer.

Procedures

A baseline examination on the 50 States was administered to students. The examiner read the instructions aloud to students, directing them to spell the States as best they could and complete the fill-in-the-blank assessment within the 20 minutes allotted. Scores were tallied and recorded in order to account for students' prior knowledge of the 50 states.

In addition to having gathered baseline data on students' 50 States knowledge, an analysis of Strand 18 scores on Spatial Relationships from the 2009 Math CMT was conducted. Matched pairs were developed using like CMT Spatial Relationships scores as the first set of criteria and U.S.A. baseline test scores as the second set of criteria. Then, each member was randomly assigned to control and experimental groups by putting both members of the pair into a container and randomly choosing one to be part of the control group. Consequently, the second to be selected each time was assigned to the experimental group.

As part of the computer-based 50 State intervention, each class met at the computer lab for three successive days. Students in the control group were directed to one side of the room and members of the experimental group were directed to the other side. Students opened an electronic version of the directions for the task. Those in the control group used a link to access Sheppardsoftware's States Level 3 game whereas students in the experimental

group had links that were set up for States Level 6. Links were masked to ensure participants played the correct game and to deter them from easily accessing the publicly available website outside of the study. Students were directed to play 3 consecutive trials of the game, then stop, record and save their scores. All students had prior training on how to take screenshots and save work to network file folders. Students repeated this process of practicing 3 consecutive trials for the following two days. Scores saved in the network folder were recorded. When browser errors or student absences occurred, those participants affected were noted. In total, 24 students (12 pairs) were affected by errors or absences and dropped from the study.

Upon finishing the second day of trials, a writing prompt was administered to students to ascertain their rating of the computer-based games. Fifteen minutes were provided for each student to complete the prompt. In their answers, students were asked to rate their interest in the software program in which they participated (Level 3 or Level 6) by taking a position that the game was either "very fun", "just all right" or "not very fun at all." When finished with the prompt, the administrator asked each student whether or not he or she had played the game at any time outside the in-class practice sessions and then recorded each student's answer.

Results

The purpose of this investigation was to examine the effectiveness of computer-based geography games as an instructional tool for teaching the 50 states of the U.S.A. Furthermore, we were interested in testing the correlation between CMT scores and the initial

baseline U.S.A. test as predictors of student success when playing the computer-based games. To accomplish this, data from student CMT scores and baseline 50 states tests were used to create matched pairs. Students in the control group played the States Level 3 game for 3 consecutive days, completing 3 trials each day. Students in the experimental group played States Level 6 following the same procedure. Data was collected on students' overall percentage correct, amount of error in distance and time needed to complete the game. Statistical modeling, including t-test and regression analyses, were used to determine the level of effectiveness and the correlation between both CMT scores and initial baseline U.S.A tests as predictors of student performance on the computer-based games.

To determine the effect which playing the game for 9 trials over the course of three consecutive days had on students' scores, a t-test was conducted to determine statistical significance by comparing scores from the first trial with the last trial. As shown in Table 2, the Sig. (2 tailed) is .000 when rounded to the nearest thousandth for each of the listed data sets. The Sig. (2 tailed) score of less than .05 means that a statistical level of significance existed between initial and final scores for percent correct, distance error in miles and number of seconds needed to play for both the control and experimental groups. Therefore, one may reject the null hypothesis that repeatedly playing the games had no effect on student performance.

Table 2
T-Test: Comparison of Student Performance Between First and Last Trials

| | | | | 95% Confidence Interval | |
| | | | | | |
Group	Mean	Std. Deviation	Std. Error Mean	Lower	Upper
Percent Correct					
Control	-17.750	9.251	1.463	-20.709	-14.791
Experimental	-24.400	9.546	1.509	-27.453	-21.347
Distance Error (mi.)					
Control	139.075	110.547	17.479	103.720	174.430
Experimental	458.900	310.307	49.064	359.659	558.141
Seconds					

Control	207.400	94.646	14.965	177.131	237.669
Experimental	244.150	156.377	24.725	194.138	294.162

Table 2 (Continued)
T-Test: Comparison of Student Performance Between First and Last Trials

	Paired Differences		
Group	t	df	Sig. (2-tailed)
Percent Correct			
Control	-12.135	39	.000
Experimental	-16.167	39	.000
Distance Error (mi.)			
Control	7.957	39	.000
Experimental	9.353	39	.000
Seconds			
Control	13.859	39	.000
Experimental	9.874	39	.000

The data collected also revealed that after completing the series of trials, all groups of students scored a higher percentage correct, improved their distance error in miles and finished the games more quickly. As shown in Table 3, the control group improved their mean score for percent correct by 17.75 percentage points and the experimental group increased by 24.40 percentage points. For mean distance error, the control group showed a reduction 139.08 miles while the experimental decreased by 458.90 miles. Additionally, the game time in seconds for the control group fell by 207.40 (3 minutes, 27 seconds) and experimental times decreased by 244.15 (4 minutes, 4 seconds).

Table 3
Comparison of Mean Scores

Group	First Trial	Last Trial	Difference
Percent Correct			
Control	69.90	87.65	17.75
Experimental	58.15	82.55	24.40
Distance Error (mi.)			
Control	178.98	39.90	-139.08
Experimental	570.73	111.83	-458.90
Seconds			
Control	445.25	237.85	-207.40
Experimental	530.48	286.33	-244.15

Similar to the analysis presented in Table 3, we compared mean scores for the sub-group of 11 participants identified with disabilities. The data revealed that students with disabilities in both the control and experimental groups increased their percentage of correct responses with the experimental group experiencing a 28.33 percentage increase and the control group 15.60 percentage increase. For mean distance error, both the control and the experimental groups decreased. Additionally, the game time in seconds for the control group decreased by 189.4 seconds (3 minutes, 4.4 seconds) and experimental times decreased by 445.83 (7 minutes, 25.83 seconds).

Table 4

Comparison of Mean Scores for Participants with Disabilities

Group	First Trial	Last Trial	Difference
Percent Correct			
Control	64.80	80.40	15.60
Experimental	50.00	78.33	28.33
Distance Error (mi.)			
Control	258.60	72.60	-186.00
Experimental	832.33	198.17	-634.16
Seconds			
Control	394.80	205.40	-189.40
Experimental	782.83	337.00	-445.83

In addition to the t-test, an SPSS regression analysis was also completed in order to test the level of correlation between both CMT scores and initial baseline U.S.A tests as predictors of student performance on the computer-based games. With CMT scores as the dependent variable, percent correct had an r value of 0.292. Therefore, CMT scores are a weak predictor of student outcomes for percent correct. However, the r value of percent correct for U.S.A. tests as dependent variable was 0.627, meaning for percent correct, U.S.A. tests were a strong predictor of student performance. As was the case with percent correct, CMT scores also showed weak r values for distance error and seconds at 0.167 and 0.170, respectively. In comparison, the U.S.A. tests proved a strong predictor of student outcomes in distance error with an r value of 0.549, but had low to medium correlation in terms of seconds needed to play with an r value of 0.369.

According to the data in Table 5, students' responses to the writing prompt indicated they liked playing the game more than they disliked it with an overall weighted mean of 2.34. The data revealed students in the control group liked States Level 3 slightly better with 0.00% students responding that it was not very fun at all versus 7.50% from the experimental group who thought States Level 6 was not very fun at all.

Table 5
Students' Rating of Computer-based Mapping Software Games

Group	Very Fun	Just All Right	Not Very Fun at All	Weighted Mean
Overall	37.50 %	58.75%	3.75%	2.34
Control	40.00%	60.00%	0.00%	2.40
Experimental	35.00%	57.50%	7.50%	2.28

Limitations

There were several threats to internal validity which arose throughout the completion of this project. These limitations included experiment effects, subject effects and history effects.

Experiment effects have been identified as a possible limitation of this study. Experiment effects arise when participants are affected by the experimenter's characteristics or expectations. In this case, the author's role as study designer, implementer, and data analyst may have skewed the data due to personal biases and affected the results of this study.

Subject effects are another possible limitation of this study. Subject effects can occur when subjects are aware of their participation in a study. Students who participated in this study had knowledge that data based on their performance would be used, in part, by the author as part of his graduate coursework. In addition, most participants recognized they had been assigned to different groups, possibly meaning they may have seen themselves in competition with students from another group. Subject effects may have affected the results of this study.

In addition, history effects were noted as a possible limitation during the collection of data. When extraneous events happen beyond the control of the experimenter, history effects can occur. Though a link was created to mask the address of the software program from students at school, it is a publicly available website which students could have accessed from home. During questioning pertaining to use of the site outside the classroom, only one participant reported searching for the program and using it outside of school; this student was removed from the study, However, there is a chance more participants may have visited the site outside of school and not reported it. Therefore, history effects could affect the results of this study.

Experiment effects, subject effects and history effects all may have had implications on the outcomes of this study. The author's role as study designer, implementer, and data analyst may have affected results through experiment effects. In addition, students' awareness of their participation in the study may have led to subject effects that could have affected results. If participants used Sheppardsoftware.com to practice the mapping game outside the classroom, history effects could also have factored into the results.

Discussion

One aim of my study was to determine the effectiveness of a computer-based software program with regards to acquisition of knowledge of the 50 states. The data collected revealed that playing the 50 state games for 9 trials over 3 consecutive days resulted in significant statistical differences in the results. These results support the findings of previous studies such as Tüzün et. al. (2009) and Carpenter and Pashler (2007) which showed computer games to be effective instructional supports when used to teach geography

Tüzün et. al. (2009) found that 4th and 5th grade geography students made significant learning gains by participating in game-based learning. In line with the results from Tüzün et. Al. (2009), participants in this study also significantly improved their mean scores. The control group playing States Level 3 improved scores by 17.75%, reduced distance error by 139.08 miles and reduced the time needed to complete the activity by 207.40 seconds, whereas the experimental group playing States Level 6 improved scores by 24.40%, reduced distance error by 458.90 miles and reduced the time needed to complete the activity by 244.15 seconds. Statistical significance was established for both the control and experimental groups by using the Sig. (2 tailed) figures of .000 when rounded to the nearest thousandth for percent correct, distance error and seconds. Though students in the Tüzün et. al. (2009) study were younger, had different learning objectives and used different games than students in this study, the results for both groups shared the implementation of computer games to produce statistically significant improvement in student achievement.

In another study, Carpenter and Pashler (2007) found that the testing effect, the process of repetitive testing which has been shown to improve performance in verbal memory tasks, helped those who were taught maps through repetitive testing make significantly larger gains than those who had extra opportunities and extended time to restudy the material. Carpenter and Pashler (2007) demonstrated statistical significance by using a 2 × 4 mixed Analysis of Variance to evaluate four separate scoring procedures. Across the four scoring procedures, mean statistics of the 2 × 4 mixed Analysis of Variance were $F(1, 46) = 6.73$, p < .043, $MSE = .024$. The T-Test conducted on the results from this study showed similar significant improvement of mean scores, as shown by the Sig. (2 tailed) figures of .000 for percent correct, distance error and seconds. The testing effect, as shown by Carpenter and Pashler (2007) to be an effective learning strategy for geography, seemed to play an important role in the success the students achieved over the course of 9 trials on three consecutive days. The games students repeatedly played from Sheppardsoftware.com appear to have utilized at least some of the potential effectiveness of the testing effect in the geography classroom.

According to the data collected from this study, the control group, which played States Level 3, had better mean scores than the experimental group, which played States Level 6. This pattern held on both the first and last trials for all categories recorded – percent correct, distance error and seconds. For percent correct, the control group outscored the experimental on the first trial by 11.75 (69.90, 58.15) and last trial by 5.10 (87.65, 82.55). It is

interesting to note that though the experimental group still scored lower on the last trial, they closed the gap that existed in the first trial by 6.65 (11.75, 5.10). A similar pattern also held for distance error and seconds. The control group initially had less distance error in miles by 391.75 (178.98, 570.73) and finished the last trial 71.93 miles closer (39.90, 111.83). From first to last, the gap for distance error decreased by 319.82 miles (391.75, 71.93), this representing a big improvement for the experimental group. Again, as in percent correct and distance error, the control group had better scores in seconds than the experimental group. The control group completed the first trial 85.23 seconds faster (445.25, 530.48) and the last trial 48.48 seconds faster (237.85, 286.33), with the experimental group closing the initial gap by 36.75 seconds (85.23, 48.48). In order to account for this pattern, it is possible to assume that States Level 6, played by the experimental group, was more difficult due to the lesser availability of visual cues available in States Level 3, played by the control group. Results from Friedman's (2009) study support this conclusion, as participants working on a computer software game gave better estimates for geographic locations when provided with more visual cues. Using Group × Region × Estimate Number ANOVAs, Friedman (2009) found main effects of estimate number for each measure: $(F, 1, 102) = 30.83$, $MSE = .12$, $\eta^2_p = .232$.

Another aim of this study was to learn more about which measures might provide the best predictive value for student outcomes on the computer-based mapping tasks. Regression analyses showed that initial baseline tests assessing students' level of prior knowledge of the 50 states proved a stronger predictor of student performance than spatial reasoning scores taken from CMT math tests. U.S.A. initial baseline tests proved a strong predictor of student performance in 2 of 3 categories, whereas, unexpectedly, CMT scores were weak predictors in all three.

Keehner et. al (2008) found a main effect between spatial ability and performance in their study using an ANOVA. Results were reported as $F(1, 56) = 8.42$, $p = .005$, $\eta^2_p = .13$. According to the findings of Keehner et. al (2008), one would expect spatial ability to have a strong correlation with performance. However, the results of the regression analyses completed in this study only support CMT math scores on spatial reasoning as weak predictors of student performance, as the r value for percent correct was .292 and had significance at the .009 level. Thus, in this study, CMT math scores for spatial reasoning did not translate into good predictors for student performance on geographical tasks presented by computer-based games.

In line with Keehner et. al (2008), Seufert et al. (2008) not only found that spatial ability contributed significantly to the regression in their analyses ($\beta = .22$, $t(77) = 2.13$, p < .05), but also prior knowledge ($\beta = .41$, $t(77) = 4.13$, p < .001). Furthermore, Coluccia et al. (2007) reported that ability to produce baseline sketch maps predicted performance on a series of geographic-related tasks (multiple $R = .66$, p < .001; adjusted R-squared = .42). Regression analyses from this study also support the strength of prior knowledge as a predictor for student outcomes as the r value for percent correct with U.S.A.

tests as dependent variable was .627 and had significance less than .001.

Besides examining the effectiveness of the computer games and the predictive values of CMT and U.S.A. test scores had on performance, we were also interested in students' ratings of the software. Overall, we found that students were motivated to complete the activity; the weighted mean used to score their responses on the written prompt shows a favorable rating of 2.34. It was interesting to note, however, that weighted mean scores showed the control group playing States Level 3 liked it slightly better (2.40) than the experimental group who played States Level 6 (2.48). This may have had to do with the increased level of difficulty, as evidenced by experimental group's lower mean scores in percent correct, distance error and seconds. Tüzün et. al. (2009) found that changes in students' intrinsic level of motivation were statistically significant (t (12) = 2.21, p < .05). The mean score for intrinsic motivation construct in the game context (M = 31.4, SD = 6.7) was higher than the mean score for intrinsic motivation construct in the school context (M = 27.6, SD = 7.2), suggesting students had higher motivation levels within the game context than the school context. As previously noted, computer-based geography instructional tools were not only successfully implemented with a statistical level of significance, but also rated favorably in terms of motivation for both this study and the Tüzün et. al. (2009) studies.

The results suggest there are several recommendations for educational practice that can be applied in the field of geography. The use of Sheppardsoftware.com to teach the 50 states proved an extremely effective tool

that may be used in the classroom or by students at home. Furthermore, the use of pre-testing to determine students' baseline level of knowledge proved a good predictor of success. Hence, the implementation of baseline tests throughout my curriculum is recommended. In contrast, CMT math scores did not translate well to the geography tasks performed during this study and are not recommended for application in the future.

Based upon the results of this study, there are a myriad of possible different avenues for future researchers to investigate. Though we learned that the games were effective and enjoyable for a period of 9 trials over 3 consecutive days, at what point might student levels of effectiveness and motivation plateau or decline? In addition, it would be intriguing to use a pre-test/post-test model similar to the procedure used by Onyancha et. al. (2009) comparing the performance of students who experienced computer-based instruction versus more traditional techniques. Accepting the data that CMT scores were not strong predictors of performance, it would be interesting to investigate whether there are other measures of spatial ability that might prove more effective.

The findings of this study suggest States Level 3 and States Level 6, created by Sheppardsoftware.com, are significantly effective instructional tools when used repeatedly for nine trials over the course of three consecutive days. Moreover, students reported they enjoyed playing the games. While math CMT scores in spatial reasoning did not prove strong predictors of students' performance, U.S.A. baseline testing was a strong indicator in 2 of 3 categories. Ultimately, computer-based

games such as States Level 3 and States Level 6 by Sheppardsoftware.com

should be considered for use in the social studies classroom.

References

Annetta, L. A. (2008). Video games in education: Why they should be used and how
they are being used. *Theory Into Practice, 47*(3), 229-239. doi:10.1080/00405840802153940

Armstrong, T. (2001). *Multiple intelligences in the classroom* (2nd ed). Alexandria, VA:
ASCD.

Bednarz, S. W., Acheson, G., & Bednarz, R. S. (2006). Maps and map learning in social
studies. *Social Education, 70*(7), 398-432.

Carpenter, S. K., & Pashler, H. (2007). Testing beyond words: Using tests to enhance
visuospatial map learning. *Psychonomic Bulletin & Review, 14*(3), 474-478.

Coluccia, E., Bosco, A., & Brandimonte, M. A. (2007). The role of visuo-spatial working
memory in map learning: New findings from a map drawing paradigm.
Psychological Research, 71(3), 359-372. doi:10.1007/s00426-006-0090-2

Committee on the Support for Thinking Spatially: Incorportion of Geographic
Information Science Across the K-12 Curriculum, Committee on Geography,
National Research Council. (2005). *Learning to think spatially: GIS as a support
system in the k-12 curriculum.* Washington D. C.: National Academy Press.

Friedman, A. (2009). The role of categories and spatial cuing in global-scale location
estimates. *Journal of Experimental Psychology / Learning, Memory & Cognition,
35*(1), 94-112. doi:10.1037/a0013590

Gersmehl, P. J., & Gersmehl, C. A. (2007). Spatial thinking by young children:
Nuerologic evidence for early development and "Educability". *Journal of
Geography, 106*(5), 181-191. doi:10.1080/00221340701809108

Hong, J. C., Cheng, C. L., Hwang, M. Y., Lee, C. K., & Chang, H. Y. (2009). Assessing
the educational values of digital games. *Journal of Computer Assisted Learning,
25*(5), 423-437. doi:10.1111/j.1365-2729.2009.00319

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences,
learning and instruction.* Hilldale, NJ: Lawrence Erlbaum Associates, Inc.

Kebritchi, M., Hirumi, A., Kappers, W., & Henry, R. (2009). Analysis of the supporting
websites for the use of instructional games in K-12 settings. *British Journal of
Educational Technology, 40*(4), 733-754. doi:10.1111/j.1467-8535.2008.00854

Keehner, M., Hegarty, M., Cohen, C., Khooshabeh, P., & Montello, D. R. (2008). Spatial

reasoning with external visualizations: What matters is what yousee, not whether you interact. *Cognitive Science, 32*(7), 1099-1132. doi:10.1080/03640210801898177

Liu, L. & Johnson, D. (2005). Web-based resources and applications. *Computers in the Schools*, *21*(3), 131-147. doi:10.1300/J025v21n03_13

Moreno-Ger, P., Burgos, D., & Torrente, J. (2009). Digital games in eLearning environments: Current uses and emerging trends. *Simulation & Gaming, 40*(5), 669-687. doi:10.1177/1046878109340294

Onyancha, R. M., Derov, M., & Kinsey, B. L. (2009). Improvements in spatial ability as a result of targeted training and computer-aided design software use: Analyses of object geometries and rotation types. *Journal of Engineering Education, 98*(2), 157-167.

Rediscovering Geography Committee, National Research Council. (1997). *Rediscovering geography: New relevance for science and society.* Washington D. C.: National Academy Press.

Salend, S. (2009). Technology-based classroom assessments: Alternatives to testing. *Teaching Exceptional Children, 41*(6), 49-58.

Sang, M., Brescia, W., & Kissinger, D. (2009). Computer Use and Academic Development in Secondary Schools. *Computers in the Schools, 26*(3), 224-235. doi:10.1080/07380560903095204

Sheppard Software, (2007). USA States – Level 3 – Online Learning. Retrieved from http://sheppardsoftware.com/USA_Geography/USA_G2_double2.html

Sheppard Software, (2007). USA States – Level 6 – Online Learning. Retrieved from http://sheppardsoftware.com/USA_Geography/USA_G5_1024_768.html

Smith, G. G., Morey, J., & Tjoe, E. (2007). Feature masking in computer game promotes visual imagery. Journal of Educational Computing Research, 36(3), 351-372.

Squire, K. D., DeVane, B., & Durga, S. (2008). Designing centers of expertise for academic learning through video games. *Theory Into Practice, 47*(3), 240-251.

Stavridou, F., & Kakana, D. (2008). Graphic abilities in relation to mathematical and scientific ability in adolescents. *Educational Research, 50*(1), 75-93.

Seufert, T., Schütze, M., & Brünken, R. (2008). Memory characteristics and modality in multimedia learning: An aptitude–treatment–interaction study. *Learning & Instruction, 19*(1), 28-42. doi:10.1016/j.learninstruc.2008.01.002

Tüzün, H., Yilmaz-Soylu, M., Karakus, T., Inal, Y., & Kizilkaya, G. (2009). The effects of computer games on primary school students' achievement and motivation in geography learning. *Computers & Education, 52*(1), 68-77. doi:10.1016/j.compedu.2008.06.008