

Effects of game technology on elementary student learning in mathematics

Namsoo Shin, LeeAnn M. Sutherland, Cathleen A. Norris and Elliot Soloway

Namsoo Shin is a research scientist in the School of Education at the University of Michigan. Her interests are focused on the impact of constructivist learning environments on learning, especially everyday problem-solving skills. LeeAnn M. Sutherland is a research scientist in the School of Education at the University of Michigan. Cathleen A. Norris is a professor in the Department of Technology and Cognition at the University of North Texas. Elliot Soloway is an Arthur F. Thurnau Professor in the Department of EECS, College of Engineering, School of Education, and School of Information, at the University of Michigan. Address for correspondence: Dr Namsoo Shin, 4035 SEB, 610 E. University, School of Education, University of Michigan, Ann Arbor, MI 48109-1259, USA. Tel: +1-734-647-4253; fax: +1-734-615-5245; email: namsoo@umich.edu

Abstract

This paper reports the effects of game technology on student learning in mathematics as investigated in two data sets collected from slightly different subjects. In the first, 41 second graders (7 or 8 years old) from two classes used either a technology-based game or a paper-based game for 5 weeks. For the next 13 weeks, both classes used a technology-based game either two times per week, or more than three times per week. A quasi-experimental control-group design with repeated measures analysis of variance and analysis of covariance was employed to explore performance differences between groups. The second study examined student learning in relation to characteristics such as their game performance, attitudes toward the game and toward mathematics, and gender and ethnicity. During a 4-month period, 50 second grade students from three classes played a technology-based game under conditions that varied depending on their teacher's direction. Multiple regression was used to determine the relationship between students' arithmetic scores and learner characteristics. Descriptive analyses by ability level, gender and ethnicity, and interview data about attitudes toward the technology game were also analyzed. Results from the two studies revealed that using a technology-based game in the classroom was beneficial to students of all ability levels in learning arithmetic skills.

Practitioner notes

What is already known about this topic

- Technology and games have yielded consistently positive results with regard to motivation, persistence, curiosity, attention and attitude toward learning.
- The accessibility and flexibility of handheld devices can provide students with more integral and spontaneous opportunities to fully take advantage of “off-moments” from classroom activity to support learning.
- There are insufficient data to validate the contribution of technology to student learning.

What this paper adds

- This paper provides evidence that game technology positively impacts elementary students' learning of arithmetic, regardless of ability level.
- The paper proposes research-based design principles, generated from a literature review, to guide the development of individualized technology game environments to support student learning.
- This paper suggests various research methodologies to provide empirical evidence of the impact of technology on academic achievement.

Implications for practice and/or policy

- To select appropriate games for target students, practitioners should consider the importance of clear goals and rules, flexible learner control and tasks at an appropriate level of challenge for the learner, with feedback provided.
- Teachers are more likely to implement technology if learning goals meet state or federally mandated standards for which teachers are already accountable.
- Identifying significant classroom variables (eg. teacher, student and classroom variables) is an important step in successfully implementing technology in educational settings. Further research is needed to understand how such variables influence learners' engagement in technology-based games and which circumstances are most advantageous for successfully implementing technology-based games in classrooms.

Introduction

For decades, technology has been used to support student learning in K-12 education. Yet, despite studies that show the potential of technology to promote student learning, many still contend that insufficient data validate the contributions of technology to student learning. For example, research using meta-analysis on the impact of technology on reading and mathematics reported that only a small number of rigorous, published studies contrast traditional and technology learning conditions for K-12 students (Dynarski *et al.*, 2007). The results of most studies in game technology were considered too fragmented and unsystematic (Condie & Munro, 2007; Randel, Morris, Wetzel, & Whitehill, 1992). The evidence is insufficient that using technology positively affected teaching and learning. Thus, the call continues for more and better empirical evidence of the impact of technology on academic achievement (Chandra & Lloyd, 2008; Dynarski *et al.*, 2007; Means, Toyama, Murphy, Bakia & Jones, 2009).

In response to the call for additional empiric evidence on the use of technology in teaching and learning, we designed two studies to investigate whether game technology facilitates learning in elementary mathematics and to explore which aspects of game design support student learning. We first reviewed research literature that explores how game technology influences student learning. Next, we developed a technology-based game using design principles generated from the review. We then conducted two quantitative studies using the technology-based arithmetic game and a paper-based game, both with the same learning goals, in order to explore the effects of game technology on student learning. This paper begins with a literature review focused on the relationship between game technology and learning, and uses that framework to propose design principles for the development of technology-based games.

Game technology and learning

Learning refers to the active construction of an individual's own knowledge by integrating new information with previous experience (Bransford, Brown & Cocking, 2000). Researchers in the learning progression field propose that because learners have different personal and cultural

experiences, they can follow many paths as they move from novice toward expert understanding (Duschl, Schweingruber & Shouse, 2007; Shin, Stevens & Krajcik, 2010). Thus, given the same learning goals, different students learn in different ways in a classroom environment, differing in the ways they perceive and comprehend information and requiring very different ways of approaching and understanding content (Sutherland, Shin & McCall, 2010). Therefore, the traditional medium of print is too limiting to meet the challenges of diverse learners' needs. For example, print materials present information for typical learners using one mode of representation (a visual mode such as pictures or text) without providing appropriate supports for individuals who learn best through other modes (eg. audio or video). Educational materials with multiple representations (eg. text, picture, video, animation/simulation, audio), various difficulty levels of learning tasks and different levels of support are necessary to appeal to the abilities, interests and needs of individual learners (Rose, Meyer & Hitchcock, 2005). As one way to address individualized materials, educators have integrated technology in games that enable students to manage the acquisition of new information based on their prior knowledge and experience, learning progress, learning style, preferences and needs (Condie & Munro, 2007). Research has shown that game features including clear goal/rules, learner control, challenging tasks/feedback and repetition enable educators to create individualized learning environments wherein students are able to take more active control of their learning and improve their achievement vis-à-vis desired learning goals (Condie & Munro, 2007; Dempsey, Haynes, Lucassen & Casey, 2002). Each of these game features and the research supporting it are described below.

Goal/rules

Games should guide learners by defining clear learning goals and specific rules. The rules of a game specify how a learner is to accomplish game goals within the limited game environment (Garris, Ahlers & Driskell, 2002). Understanding goals and rules supports student cognitive development, especially organizational skills and abstract thinking. Learners need to develop a mental structure to understand the underlying concepts of a game including its goals/rules, procedures, properties and conditions (Garris *et al.*, 2002). The development of such a mental structure facilitates the development of organizational skills (Vygotsky, 1978). In addition, playing games enhances abstract thinking such as organizing ideas with particular properties into groups and creating sophisticated generalizations such as finding patterns (Piaget & Inhelder, 1967). Technology-based games make extensive use of symbols and other abstractions; therefore, to be successful, the learner needs to understand the symbolic meaning of objects in a game and use the symbols and general strategies to form hypotheses and to experiment. Vygotsky believed that performing actions according to rules and symbolic understanding is a valuable experience for children, as the process of understanding symbolic meaning is considered the first step toward abstract thinking (1978). Research has shown that the goals and rules adopted by learners during a learning task influence both their performance and their motivation (Locke & Latham, 1990; Urdan & Maehr, 1995; Wolters, Yu & Pintrich, 1996). These findings support the notion that those educational games that contain clear and specific goals and rules lead to enhanced performance and learning outcomes.

Learner control

Games should allow learners to control their own activity and to direct or command elements of the game. Learner control is defined as the ability of learners to regulate learning activities based on their learning styles, strategies, and previous experience and knowledge (Quintana, Shin, Norris & Soloway, 2006). Vygotsky proposed that individuals construct a variety of internal developmental processes when they are interacting with others in a learning environment (1978). The functionality of learner control in games facilitates an interaction between student and technology, which is the important factor for creating an individualized learning environment to

promote student learning (Kahveci & Imamoglu, 2007; Wagner, 1994). For example, students might make decisions about the topics and difficulty level of the activity based on their ability level, task goals and game strategies (Carter, 1993; Kinzie & Joseph, 2008). Having control is an important aspect of subject matter mastery, motivation and positive attitude toward learning activities (Blumenfeld, Kempler & Krajcik, 2006). Research on learner control has yielded consistently positive results with regard to motivation and learning (Cordova & Lepper, 1996; Eck, 2006; Morrison, Ross & Baldwin, 1992).

Challenge tasks/feedback

Vygotsky advocated the notion of the zone of proximal development (ZPD) in the learning process, suggesting that learning is facilitated when individuals face activities that lie just outside of their ability (1978). On the basis of this principle, games need to offer tasks that are more challenging than students' previous accomplishments. When students work on challenging tasks using game technology, their motivation to compete against and improve their own previous scores increases (Inal & Cagiltay, 2007). Along with task challenge, games should also provide feedback (eg. embedded expert guides) to support the learner in reaching goals. A crucial point in feedback is a game's response to the actions of a learner in an intelligent way in order to encourage the learner and to guide his or her behavior toward an optimal learning goal (Wagner, 1994). Games might provide feedback based on the student's correct or incorrect individual answers, on the number of correctly solved problems out of the total score or on other factors that enable the student to either pass or fail to move on to the next level. Such feedback helps learners to evaluate whether their current performance meets established goals and to reflect on past performance (Whitehall & McDonald, 1993). Learners can then seek to improve their performance in order to reduce the discrepancy between goals and performance (Kernan & Lord, 1990). This learning process allows learners to keep track of their progress toward desired goals and leads to increased effort and performance.

Repetition

Games should call for repetitive play and repeated return to the gaming activity across time. Repetition allows learners to evaluate their performance based on trial-and-error strategies and encourages them to persist or to intensify their effort to complete required tasks (Thomas & Macredie, 1994). Research indicates that these repetitive activities influence greater achievement (Garris *et al.*, 2002).

Research literature supports the notion that game technology promotes student learning of important ideas and skills, positive motivation, persistence, curiosity, attention and attitude toward learning (Becker, 2007; Rieber, 1996; Vogel *et al.*, 2006; Whitehall & McDonald, 1993). Moreover, research suggests that game technology improves student performance on algebra and mathematics problem solving (McFarlane, Sparowhawk & Heald, 2002), reading comprehension, spelling and decoding of grammar (Rosas *et al.*, 2002), and complex thinking skills including problem solving, strategic planning and self-regulation (Cordova & Lepper, 1996; Ricci, Salas & Cannon-Bowers, 1996). Although there is potential in the use of game technology as an individualized learning tool, research findings with regard to the effectiveness of educational game technology in classroom learning are mixed (Condie & Munro, 2007; Vogel *et al.*, 2006); limited robust research exists on the use of technology-based games and the application of gaming skills and techniques in educational contexts. Further research is necessary to provide evidence of the effects of game technology on student learning in classrooms.

Purpose of the study

The primary purpose of this research was to investigate the effects of game technology on student learning of mathematics in the elementary classroom. Two studies were designed to investigate the following questions:

- How does student performance on an arithmetic skills assessment vary for students using a technology-based game and students using a non-technology-based game?
- How does student performance on an arithmetic skills assessment differ depending on whether they play the technology-based game twice or play more than three times per week?
- What is the relationship between students' performance on a technology-based arithmetic game (high/low) and their performance on a related skills assessment?

Method and results

Materials

Research has determined that to learn advanced mathematics (eg. solving problems, understanding other concepts in mathematics), students must, with quality instruction, master basic arithmetic skills (such as addition and subtraction) in the early stages of the learning experience (Bloom, 1956; Hoon, Chong & Binti Ngah, 2010). Research has also shown that game technology may be an optimal teaching and learning approach to facilitate student learning of basic mathematics skills such as arithmetic (Christmann & Badgett, 2003; Condie & Munro, 2007; Miller & Robertson, 2010).

Technology-based game: GameBoy™ Skills Arena

An educational game program called *Skills Arena* (GoKnow Learning, Dallas, TX, USA) was developed to teach students about basic arithmetic skills using the GameBoy™ (Nintendo Co., Kyoto, Japan) system. GameBoy may well enable the provision of individualized instruction because its size makes it easy for children to handle (Shin, Norris & Soloway, 2006). The accessibility and flexibility of GameBoy can provide students with more integral and spontaneous opportunities to fully take advantage of “off-moments” from classroom activity to support learning (Finn & Vandenharn, 2004). In designing the program, we used ideas from the literature about the importance of clear goals and rules; flexible learner control; tasks at an appropriate level of challenge, with feedback; and repetition (Garris *et al.*, 2002; Quintana *et al.*, 2006; Thomas & Macredie, 1994; Wagner, 1994). We specified learning goals of the game program that directly linked to curriculum objectives and specific concepts or skills. Researchers propose that teachers are more likely to implement technology if learning goals meet state or federally mandated standards for which teachers are already accountable (Norris, Shin & Soloway, 2007).

In *Skills Arena*, students create their own characters and identities (ie. name). We developed several features according to the design principles of having appropriate levels of challenge tasks for diverse learners and having flexible learner control. For example, students are able to select a game type from among three choices: “Mental,” “Extension” and “Place” games. The Mental game is the easiest, dealing only with single digit numbers such as “ $_ + 3 = 10$.” The Extension game is moderately difficult, incorporating tasks such as “ $300 + 400 =$ ” or “ $260 - 50 =$.” The Place game is the highest level. It employs multiple-digit numbers without zeros such as “ $245 + 563 =$ ” or “ $437 - 36 =$.” Each game has three options for arithmetic tasks: addition, subtraction, and mixed addition and subtraction. The arithmetic problems in *Skills Arena* are composed of one-, two-, three- and four-digit numbers. Students need to solve the problems before the questions fade out on the GameBoy screen. In addition to game type and arithmetic task, the game has options for six speed levels, presented in the form of cartoon characters. As students answer a question, they receive immediate feedback as to whether they have answered correctly or incorrectly. The game displays a performance summary immediately after each game ends. The summary shows the total number of problems attempted and the number correctly solved. It also tracks and displays the total number attempted and correctly solved for the individual over time (see Figure 1a). This feature was intended to provide appropriate feedback to students in an intelligent way based on their performance on game tasks.

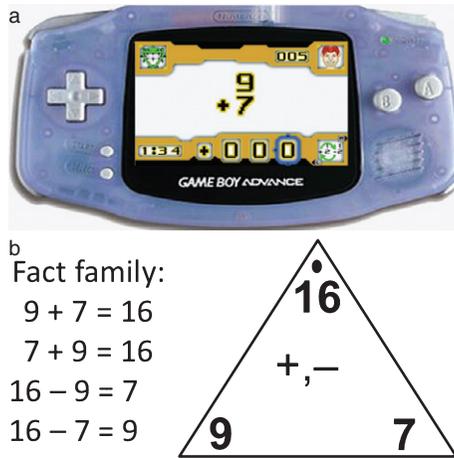


Figure 1: (a) Skills Arena game screen. (b) Flash fact triangle card

Paper-based game: flash fact cards

The flash fact card is a triangle-shaped card with three numbers, one on each angle, and both an addition and subtraction sign in the middle of the card (see Figure 1b). A pair of students is teamed as partners. One student holds a card by hiding one of the three numbers and asks the other student what number is hidden on the card. The other student in the pair answers the question using the two remaining numbers and one of the arithmetic signs (plus or minus) on the card. The teacher directs students to practice either addition or subtraction before students begin to play. Students repeat the procedure with all the cards that the questioner has and then switch roles.

The essential cognitive processes of the both the paper-based games and the technology-based games are similar in that a student needs to calculate a correct answer using two numbers and an arithmetic sign without any support. The target school had already been using the flash fact cards to teach arithmetic skills for second grade students. Given that the cognitive process, the learning goals and the target students of the *Skills Arena*, GameBoy is identical to those of the flash fact cards and that both are considered learning games, we believed it valid to compare the two methods in terms of their ability to support the learning of arithmetic. Another reason for comparing the two methods is that the school principal and teachers perceived the technology-based game to be better for practicing arithmetic skills and wanted to use the GameBoy for all second grade students. This study was designed to determine whether it is, in fact, the case that the technology-based game improves student learning.

Study One: comparison between use and non-use, and more and less technology

This study has used a quasi-experimental control-group design with repeated measures to investigate the first two research questions: (1) how do students' performances vary between a technology-based and a paper-based game during a 5-week period? and (2) how do students' performances differ when they play the game two times per week compared with playing more than three times per week during a 13-week period? During the 5-week period, the study provided information regarding the performance differences between the technology-based and non-technology-based game groups. The 13-week period was intended to explore the performance differences in relation to the frequency of playing the technology-based game per week. The rationale of the 13-week research design is to investigate the appropriate amount of playing time to efficiently maximize the effects of game technology on student learning.

Participants

The participants in Study One were 41 second grade students (7 or 8 years old) from two classes (out of three classes of second grade students in the school) attending a public elementary school near a large city in the Midwest. The third class is excluded in the research because the class composition and size were markedly different from the other two classes, in that it combined first ($n = 9$) and second ($n = 9$) grade students, and the sample size of nine second graders may not be enough to conduct statistical analysis for comparing their performances with those of other two classes for a 4-month period. Students were primarily upper-middle class and included 24 male and 17 female students. School records indicate that 28 students were white, seven were Asian, two were African-American and four were "other."

Study conditions and procedures

Because one research question focused on determining the value added of a technology-based game, it was important that both the GameBoy (GB) group and the card and GameBoy (CG) group played the arithmetic game—whether technology based or paper based—for the same amount of time. The students in the GB used the *Skills Arena* software program. Before administering the pretest, we had students practice using the GameBoy game for 15 minutes each day for 10 days. Two investigators conducted a 1-hour training session for teachers and students on how to use the GameBoy game. Although most students had previous experience playing games on the GameBoy, some of them required additional guidelines for playing the arithmetic game. After that, students played *Skills Arena* on the GameBoy for 15 minutes in their mathematics class, three times per week for 5 weeks. After 5 weeks, the students played the game for 15 minutes, twice per week for 13 weeks (see Table 1). The teacher instructed students to play the game after she had finished the regular lesson. The students were kept from playing the game except for the designated period.

The students in the CG group played the addition and subtraction *flash fact cards game* for 15 minutes in their mathematics class, three times per week. The teacher instructed students to play the *flash fact cards game* after she finished the regular lesson. After 5 weeks, the students played the GameBoy game for the remaining 13 weeks of the semester. The first 10 days with GameBoy, students in the CG group played the game the same as those in the GB group. Following that, they played the game for 15 minutes, more than three times per week, in essence to enable them to "catch up" in the amount of time they used the technology-based game. Furthermore, students were allowed to play the game anytime they finished their other class assignments (see Table 1). The teachers discussed the two conditions and decided which option they would implement in their own classrooms. GB was composed of the class of 20 students. CG consisted of the class of

Table 1: Study One research design

Procedure	GB ($n = 20$)	CG ($n = 21$)
Training	GameBoy training: 150 minutes	Card game training: 150 minutes
Pretest	X	X
5-week game period	GameBoy playing: 225 minutes (three times per week)	Card game playing: 225 minutes (three times per week)
5-week test	X	X
13-week game period	GameBoy playing: 390 minutes (two times per week)	GameBoy training: 150 minutes GameBoy playing: 585 and more (more than three times per week)
Final test	X	X

GB, GameBoy game group; CG, Card and GameBoy game group.

21 students. The pretest was administered under the supervision of teachers and two investigators before instruction began. After each period of gaming (the initial 5-week period and the second 13-week period), teachers and investigators administered the 5-week test and the final test in the two classes, respectively (see Table 1). During the implementation stage of the technology-based game, a researcher and an assistant conducted classroom observations of the GB and CG classes twice per week. We did not develop an instrument for observing teaching practices because the scope of this study did not include evaluating participants' teaching of the mathematics class. Instead, during the 5-weeks of instruction, the classroom observation focused on ensuring that the *Skills Arena* game was implemented in GB as intended. For CG, the observation was to guarantee that students' arithmetic skill practice hours would be the same as in GB. Observations revealed that there was no difference in time devoted to arithmetic skill practice between CG and GB. During the 13-week period, the classroom observation was conducted to confirm that the GB students played the game consistently in an appropriate way, and the *Skills Arena* game was precisely implemented in CG. Our observation verified that the technology-based game was implemented in CG and GB as intended.

Instruments

A 70-item instrument was developed to assess arithmetic skills as a learning outcome in this study (see Table 2). The test had two sections. Part I consisted of 50 questions with 25 addition items and 25 subtraction items to measure basic arithmetic skills. Part II of the test had 20 questions, including 10 addition and 10 subtraction items, designed to assess advanced arithmetic skills. The students took Part I of the test in 15 minutes and Part II in 5 minutes. Participating teachers confirmed that the instruments were aligned with the mathematics standards for second grade students.

Data analysis

Two male students from GB and one male student from CG were excluded from the analysis because they did not take one or more of the assessments: the pretest, the five-week test or the final test. One female student from GB was discounted because she transferred to another school. Thirty-seven students from the two groups were included in this analysis. This was a quasi-experimental control-group design with repeated measures. Repeated measures analysis of variance (ANOVA) was employed to analyze student data, as appropriate when all students in groups are measured multiple times under different conditions, and the measurement of the dependent variable is repeated (Vonesh & Chinchilli, 1997). The repeated measures ANOVA allowed researchers to monitor how students in the two groups' performance changed over three measurements during a 4-month period. Analysis of covariance (ANCOVA) was conducted, treating students' prior knowledge as a covariate variable, enabled researchers to compare the scores of two groups on the 5-week and the 13-week assessments. Effect sizes obtained from the two groups

Table 2: Sample test items

Addition	Subtraction
Part I (Basic level)	
$60 + 30 = \underline{\quad}$	$50 - 40 = \underline{\quad}$
$27 + \underline{\quad} = 150$	$24 - \underline{\quad} = 11$
$255 + 110 = \underline{\quad}$	$425 - 220 = \underline{\quad}$
Part II (Advanced level)	
$346 + 65 = \underline{\quad}$	$402 - 14 = \underline{\quad}$
$123 + 188 = \underline{\quad}$	$151 - 78 = \underline{\quad}$
$658 + 368 = \underline{\quad}$	$184 - 135 = \underline{\quad}$

(by calculating the difference between adjusted means with pooled adjusted mean standard deviations) provided additional information about student performance. Assumptions for the ANCOVA statistical measures were checked using Levene's test for equality of variances between the groups. The homogeneity-of-regression (slopes) assumption of the relationship between the covariate and the dependent variable was also examined. The data for the all students as well as for the lower level students as a subgroup met the assumptions.

The analysis for lower level student performance was conducted using repeated measures ANOVA and ANCOVA with only lower level students from both of the groups. The lower level group consists of the students who had scores lower than 42 on the pretest. Although the mean of pretest score of the both groups was around 37, a close look at frequencies revealed two distinct clusters of higher and lower level students. The frequency of the pretest scores with all students was distributed evenly across 12 and 57 except between 22 and 28, and 41 and 46. No student scored between those two score ranges. The 42 cutoff point was appropriate for determining higher and lower level groupings because if we were to use 22 as the cutoff point, there were insufficient students (less than 30%) to conduct statistical analysis by ability level. Determining high and low ability using 42 as the score around which actual scores clustered meant that 50% of the students could be classified as lower performing and 50% classified as higher performing. However, the sample size of each low-level group ($n = 10$ from GB and $n = 11$ from CG) using 42 was relatively small. Assumptions for the statistical measures were checked using Levene's test for equality of variances among the groups. The data for the low-level samples met the assumptions. The statistical analysis was interpreted to examine the effects of the technology-based game on low-level students.

Results and discussion

Descriptive statistics for the two groups are shown in Table 3. Overall, the scores of the male students in the both groups were higher than those of females across three tests. White and Asian students performed better than African-American students.

Table 4 summarizes the results of the repeated measures ANOVA analysis. The column labeled *F* gives the *F* value of the test followed by three columns of significance values. A repeated measures ANOVA with a Greenhouse–Geisser correction determined that mean of total arithmetic test scores differed between the three time points (pre, 5-week and final test) at statistically significant levels: $F(2, 70) = 33.38, p = 0.000$. There was a significant interaction between the three time points and the groups (GB and CG), $F(2, 70) = 4.57, p = 0.014$. The statistical interaction results indicated that students' performances depended on their use of the game technology.

As shown in Figure 2a, the direction of a relation between two groups is different depending on the use of game technology. When students used or played the technology-based game more, their performance was higher than that of students who did not use the technology-based game or who played it less often.

In the 5-week period, the GB students gained 11% ($M = 7.65$, standard deviation [*SD*] = 7.84) from the pretest to the 5-week test. In the case of CG, students increased only 4% ($M = 2.85$, $SD = 6.48$) on the tests. A one-way ANCOVA was conducted to compare the scores of two groups on the 5-week test. The independent variable, group level, included two levels: CG and GB. The dependent variable was the students' arithmetic scores on the 5-week test, and the covariate was the students' scores on the pretest. An analysis evaluating the homogeneity-of-regression (slopes) assumption indicated that the relationship between the covariate and the dependent variable did not differ significantly as a function of the independent variable, $F(1, 33) = 2.96, p = 0.095$. The underlying assumption of homogeneity of variance for the one-way ANCOVA has been met—as evidenced by $F(1, 35) = 0.29, p = 0.596$. The ANCOVA was significant, $F(1, 34) = 4.24, p = 0.047$ (see Table 5). A follow-up test using the Bonferroni procedure was conducted to

Table 3: Mean scores by group on each of three test administrations

	Pretest		5-week test		Final test	
	Mean (%)	SD	Mean (%)	SD	Mean (%)	SD
GB (<i>n</i> = 17)						
Test total (70)	37.06 (53)	14.31	44.71 (64)	12.62	45.53 (65)	15.22
Basic (50)	35.12 (70)	12.32	40.18 (80)	9.75	40.06 (80)	11.57
Advanced (20)	1.94 (10)	2.59	4.53 (23)	3.59	5.47 (27)	5.21
Ability						
Low (<i>n</i> = 10)	27.00 (38)	9.32	38.50 (55)	12.76	38.80 (55)	16.22
High (<i>n</i> = 7)	51.43 (74)	2.51	53.57 (77)	4.93	55.14 (79)	6.31
Gender						
Male (<i>n</i> = 9)	39.22 (56)	14.88	43.33 (63)	13.73	44.67 (64)	17.05
Female (<i>n</i> = 8)	34.63 (48)	14.22	46.25 (65)	11.97	46.50 (66)	13.97
Ethnicity						
White (<i>n</i> = 12))	40.58 (57)	3.70	47.08 (67)	3.47	48.25 (69)	3.97
African-American (<i>n</i> = 1)	22.00 (31)	—	22.00 (31)	—	24.00 (34)	—
Asian (<i>n</i> = 2)	43.00 (61)	11.00	50.50 (72)	8.50	58.50 (84)	0.50
Other (<i>n</i> = 2)	17.50 (25)	2.50	36.00 (51)	2.00	27.00 (39)	4.00
Game score						
Total ¹	—	—	917	414.82	3369.59	1403.14
Correct ²	—	—	770	423.83	2790.82	1408.22
CG (<i>n</i> = 20)						
Test total (70)	37.05 (53)	15.58	39.95 (57)	17.00	47.60 (68)	16.92
Basic (50)	34.80 (70)	13.40	34.85 (70)	14.31	39.30 (79)	11.48
Advanced (20)	2.25 (11)	2.65	5.05 (25)	3.86	8.30 (42)	6.05
Ability						
Low (<i>n</i> = 11)	25.36 (36)	10.85	28.81 (41)	15.23	37.00 (53)	15.92
High (<i>n</i> = 9)	51.33 (73)	3.50	53.44 (77)	4.61	60.56 (87)	4.42
Gender						
Male (<i>n</i> = 12)	44.67 (64)	12.38	45.67 (65)	13.49	54.25 (78)	11.46
Female (<i>n</i> = 8)	25.63 (38)	13.02	31.38 (45)	19.01	37.63 (54)	19.52
Ethnicity						
White (<i>n</i> = 14)	37.85 (54)	4.19	41.14 (59)	4.45	48.07 (69)	4.60
African-American (<i>n</i> = 1)	15.00 (21)	—	15.00 (21)	—	19.00 (27)	—
Asian (<i>n</i> = 3)	47.00 (67)	5.57	53.67 (77)	2.33	59.33 (85)	2.19
Other (<i>n</i> = 2)	27.50 (39)	11.50	23.00 (33)	6.00	41.00 (57)	11.00
Game score						
Total ¹	—	—	—	—	3560.00	2850.39
Correct ²	—	—	—	—	2371.03	1677.05

¹Total number of problems attempted.

²Number of problems correctly solved.

GB, GameBoy game group; CG, Card and GameBoy game group.

evaluate differences of the adjusted means for group level. The results showed that GB students ($M = 44.70$) had significantly higher arithmetic scores, controlling for the effect of their pretest, than CG students ($M = 39.90$). The effect size for the significant adjusted mean difference was 2.06, which indicates a large effect size (Cohen, 1988) (see Table 5). In the next 13-week period, the CG students increased 11% ($M = 7.70$, $SD = 6.12$) from the 5-week test to the final test. The GB students increased only 1% ($M = 0.82$, $SD = 6.54$) on the tests. The ANCOVA results using the 5-week scores as a covariate variable showed that CG outperformed GB, $F(1, 34) = 10.07$, $p = 0.003$. The results of the follow-up test using the adjusted means for group level showed that CG students ($M = 49.77$) had significantly higher arithmetic scores, controlling for the effect of

Table 4: Repeated measures ANOVA results across three tests

Effect	MS	df	F	p	Greenhouse–Geisser	Huynh–Feldt
All students ($n = 37$)						
Time	834.11	2	33.38	0.000	0.000	0.000
Time \times Group	114.29	2	4.57	0.014	0.015	0.014
Error	24.99	70				
Low-level students ($n = 21$)						
Time	737.56	2	22.82	0.000	0.000	0.000
Time \times Group	110.77	2	3.43	0.043	0.045	0.043
Error	32.33	38				

Note. Time = pretest, 5-week test and final test; Group = GameBoy versus Card and GameBoy. ANOVA, analysis of variance; MS, mean square; df, degrees of freedom.

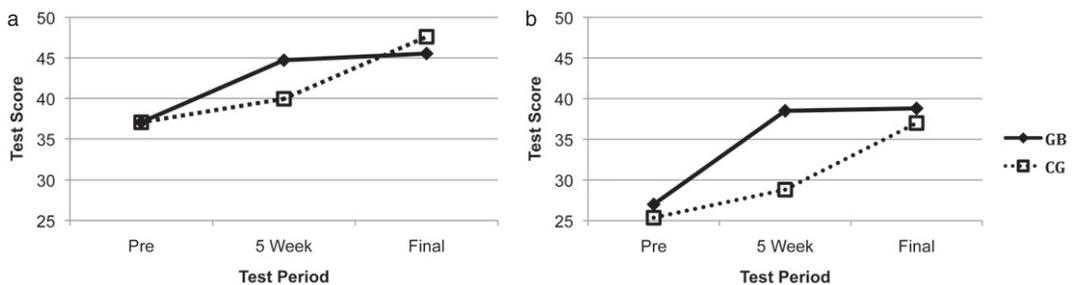


Figure 2: (a) All students across three tests by group. (b) Low-level students across three tests by group. GB, GameBoy game group; CG, card and GameBoy game group

their 5-week test, than GB students ($M = 42.98$). The effect size for the significant adjusted mean difference was 3.17, considered a large effect size (see Table 5).

The results of repeated measures ANOVA with only low-level students revealed that mean of total arithmetic test scores differed between the three time points at statistically significant levels: $F(2, 38) = 22.82, p = 0.000$ (see Table 4). There was a significant interaction between the three time points and the group, $F(2, 38) = 3.43, p = 0.043$ (see Figure 2b). The GB low-level students gained 16% ($M = 11.50, SD = 7.12$) from the pretest to the 5-week test. In case of CG, students increased only 5% ($M = 3.45, SD = 8.34$) on the tests. The ANCOVA results of the 5-week test scores showed that GB outperformed CG: $F(1, 18) = 5.16, p = 0.036$ (see Table 5). The results of the follow-up test showed that GB lower level students ($M = 37.50$) had significantly higher arithmetic scores, controlling for the effect of their pretest, than CG lower level students ($M = 29.73$). The effect size for the significant adjusted mean difference was 2.27, which was considered large. In the 13-week period, the CG low-level students increased 12% ($M = 8.18, SD = 7.32$) from the 5-week test to the final test. The GB low-level students increased approximately 0% ($M = 0.30, SD = 7.64$) on the tests. The ANCOVA results showed that there were statistically significant differences in test scores between the groups during the 13-week gaming: $F(1, 18) = 4.95, p = 0.039$ (see Table 5). The follow-up test results showed that CG low-level students ($M = 41.64$) had significantly higher arithmetic scores, controlling for the effect of their 5-week test, than GB low-level students ($M = 33.69$). The effect size for the significant adjusted mean difference was 2.22, which was large.

We ran descriptive statistics to compare student performance by gender, even though the sample size is small. In the 5-week gaming period, the GB female students' gain scores ($M = 11.62$) were twice as high as those for CG females ($M = 5.75$). In the same period, the GB males ($M = 4.1$) gained

Table 5: ANCOVA results for all students and for lower level students by group

Source	SS	df	MS	F	p
All students ($n = 37$) on 5-week test					
Pretest	6,341.39	1	6,341.39	126.98	0.000
Group	211.54	1	211.54	4.24	0.047
Error	1,697.94	34	49.94		
Total	8,251.57	36			
Pretest \times group	139.795	1	139.795	2.961	0.095
Homogeneity of variance		1		0.29	0.596
All students ($n = 37$) on 13-week test					
Five-week test	7,751.00	1	7,751.00	189.04	0.000
Group	412.70	1	412.70	10.07	0.003
Error	1,394.04	34	41.00		
Total	9,184.43	36			
Five-week test \times group	46.09	1	46.09	1.13	0.296
Homogeneity of variance		1		0.02	0.880
Low-level students ($n = 21$) on 5-week test					
Pretest	2,693.30	1	2,693.30	44.36	0.000
Group	313.417	1	313.417	5.16	0.036
Error	1,092.83	18	60.71		
Total	4,277.14	20			
Pretest \times group	0.94	1	94	0.02	0.905
Homogeneity of variance		1		0.02	0.896
Low-level students ($n = 21$) on 13-week test					
Five-week test	3,838.04	1	3,838.04	64.83	0.000
Group	292.91	1	292.91	4.95	0.039
Error	1,065.56	18	59.20		
Total	4,920.57	20			
Five-week test \times group	44.29	1	44.29	0.74	0.402
Homogeneity of variance		1		0.23	0.634

Note. Group: GameBoy versus Card and GameBoy.

ANCOVA, analysis of covariance; MS, mean square; df, degrees of freedom; SS, sum of squares.

only three points more than the CG males ($M = 1.00$). In the 13-week period, both women ($M = 6.25$) and men ($M = 8.58$) in CG gained substantially more than those in GB ($M = 0.25$ for females $M = 1.34$ for males). The trend in females' scores across the three tests indicated that low-level females' performance was positively influenced by the use of the technology-based game. GB females increased their score from $M = 34.65$ on the pretest to $M = 46.25$ on the 5-week test. Their scores on the 5-week test were higher than the cutoff point (42) dividing the low- and high-level group's scores. However, after reaching a high-level score, their scores did not increase from the 5-week to the final test. In contrast, the CG females who scored lower than 42 on the pretest and the 5-week test consistently improved their scores across the three tests. The CG males who scored more than 42 on the pretest substantially increased their scores after playing the technology-based game during the 13-week gaming period, while the GB males who had lower than 42 scores showed a trivial improvement. These results suggest that game technology might be effective in supporting low-performing female students in learning basic arithmetic skills. Results also suggest that game technology might be beneficial in supporting high-performing males in learning basic arithmetic skills. Further study is merited to generalize the results with larger and more diverse students as well as to investigate why game technology seems to more positively affect lower performing female and higher performing male students. For example, specific aspects of student attributes such as level of competitiveness and confidence in mathematics, or attitude toward group work, are necessary for analyzing student's performance by gender in depth.

In sum, the results of the scores between the pretest and the 5-week test revealed that students who used the technology-based game outperformed those who did not use it. Comparison of the 5-week and the final test results revealed that students who played the technology-based game more than three times per week outperformed those who used the game only twice each week. The trends for low-level students as a subgroup were the same as those for all students.

However, in order to reach firm conclusions about the effects of game technology on student learning, further investigation must continue to explore the results of this study. First, the 13-week results might have been influenced by other factors such as motivation toward the game and competence in using the game. For example, students in the two groups might have different degrees of motivation toward the technology-based game before starting the 13-week gaming. The GB students might have low motivation toward the game because they had already used it for 5 weeks. In contrast, CG students might have higher motivation than those of GB students because of the novelty of using the game and the *GameBoy*. We employed an ANCOVA statistical technique to control for prior performance differences between the groups but did not measure and control motivation and competence differences toward the technology-based game in the 13-week analysis.

Second, we need further investigation to explore the lack of progress of GB students from the 5-week to the final test despite 13 weeks of using the game. We speculate three possible reasons for the results: first, because the novelty of using the game wears off during the 4-month period, the GB students' motivation toward the game might decrease, resulting in a no-gain score. Second, twice weekly use of the game may sustain students' arithmetic skills, but may not be sufficient for improving their skills. Technology should be integrated seamlessly into the school curriculum as a routine learning activity for successful implementation in classrooms. However, the twice-a-week schedule of playing arithmetic games might not be sufficient for establishing the use of game technology as an integral part of the classroom learning experience. Third, students' scores on the final test might be a factor of their developmental growth rather than the effects of game technology. That is, with or without using a technology-based game, 7- or 8-year-old children might be typically able to reach only a certain arithmetic skills level developmentally (eg. correctly answer around 50 of 70 items).

Finally, the nonsignificant difference between two groups on the final test score requires additional exploration. The results are congruent with those of the final game score (see Table 3). No statistically significant differences emerged between the two groups on the cumulative number of problems solved or of problems solved correctly as indicated by the *GameBoy* game scores. These results suggest that either the number of problems attempted or number of correctly solved problems may influence students' scores on the arithmetic test regardless of frequency of playing the game (twice vs. more than three times per week), or duration in using the technology-based game (18 weeks vs. 13 weeks). Developmental growth may also help to explain the lack of significance in the two groups' results on the final test. It may be that the effects of game technology are limited in promoting drill-and-practice learning tasks (eg. arithmetic skills), in part, because students typically can only reach a certain level based on their developmental stage regardless of using game technology. We also speculate that our three tests during the 4-month period might not sufficiently discriminate students' learning progress. The 70 items might not include an appropriate range of item difficulty. For example, the 20 items in Part II might be too difficult to answer correctly within 5 minutes. The instrument might lack sufficient middle-level items for measuring student performance in the score range between 50 and 70.

To answer these questions and to confirm our speculations from the first study, additional research using a randomized control trial with control and treatment groups for a longer period time, with a well-developed test and a larger sample size is necessary. In addition, student data

regarding attitude toward and competence in using the game are needed in conjunction with administration of the arithmetic tests (eg. pretest, five-week test and final test). Such study can inform whether our test adequately discriminated the effects of game technology over a 4-month period. Such study might also provide information as to the reasons for students' gain scores, whether the gains originate from students' developmental growth or from the effects of game technology.

Study two: relationship between learner characteristics and learning using a technology-based game

The purpose of Study Two was to explore relationships between student characteristics (gender, ethnicity, attitude toward mathematics, attitude toward the game, game performance) and student learning in a game technology environment. Study Two was designed to investigate: how do students' performances on a technology-based game influence student learning of arithmetic skills? This question was investigated to determine whether particular characteristics could predict students' arithmetic scores. More specifically, if game performance can predict (at a level of statistical significance) students' scores on an arithmetic test, we can then determine that game technology influences student learning in mathematics.

Participants

Participants in this study included the 41 students from Study One as well as the nine students from class three who were omitted from Study One analysis. The third class is included in this research design because the rationale of the research design was to explore the relationship between students' game scores in *GameBoy* and student's scores in the arithmetic test regardless of their class conditions for using the game. In Study Two, we treated individual characteristics as important variables and did not intend to compare students' performances between three class gaming conditions. The 50 second graders from three classes of 21, 20 and 9 students each, included 28 male and 22 female students. School records indicated race/ethnicity of the students as follows: white ($n = 31$), African-American ($n = 6$), Asian ($n = 8$), Hispanic ($n = 1$) and other ($n = 4$) (see Table 6 for detail).

Instrument

The subjects' attitude toward mathematics was measured using the attitude toward math survey. This survey, developed by Pearce, Lungren and Wince (1999), is comprised of 23 positive and negative statements about young students' attitudes toward mathematics. We used this survey

Table 6: Participants and research design for Study Two ($n = 50$)

	Total	Class One ($n = 20$)	Class Two ($n = 21$)	Class Three ($n = 9$)
Duration of using GameBoy (Minutes of playing game)		18 weeks (615 minutes)	13 weeks (+585 minutes)	8 weeks (+480 minutes)
Frequency of playing ¹		5 weeks: 3× per week 13 weeks: 2× per week	More than 3× per week	More than 4× per week
Gender				
Male	28	11	13	4
Female	22	9	8	5
Ethnicity				
White	31	14	14	3
African-American	6	1	1	4
Asian	8	3	4	1
Hispanic	1	—	—	1
Other	4	2	2	—

¹Frequency of playing GameBoy per week.

because it was developed at an age-appropriate level for lower elementary students. The teachers presented statements to students orally. Students responded by marking “yes” or “no” on their answer sheet in response to each statement. An eight-item debriefing questionnaire was developed to capture students’ perceptions regarding attitudes toward elements of the game. The items were a combination of six “Yes or No” and two open-ended questions (see Table 7). Participating teachers confirmed that the instruments were at an appropriate readability level for second grade students. The 70-item instrument developed in Study One was used to assess arithmetic skills as a learning outcome (see Table 2).

Procedures

Fifty second grade students from three classes used a GameBoy with the *Skills Arena* software program under three conditions including three times per week for 5 weeks followed by two times per week for 13 weeks over an 18-week period, more than three times per week over a 13-week period or more than four times per week over an 8-week period (see Table 6). We designed three conditions to ensure that each group would spend approximately the same amount of time playing with the technology-based game. Thus, the primary difference among three groups was the frequency and the duration in using the technology-based game, keeping a total amount of playing time the same approximately. The students in Class One and Class Two played the game as described in Study One. The students in Class Three played the GameBoy game more than four times per week for 8 weeks. Instructions for playing the game were the same as those of GB in Study One. The teachers and investigators administered the final test in the three classes. The teachers administered the survey of students’ attitudes toward mathematics. A researcher collected students’ responses of the debriefing questionnaire through individual interviews with each participant after administrating the final test.

Data analysis

Multiple regression (a nonexperimental, correlational analysis) was used to test the relationship between students’ arithmetic scores and the predictor variables including attitude toward the game and toward mathematics, ethnicity, gender and game scores (number of correctly solved problems). Pearson correlation coefficients were computed for the entire sample in order to detect violation of the data assumptions (multicollinearity among the independent variables: correlation higher than 0.70) for conducting multiple regression (see Table 8). The mathematic attitude scores were significantly related to those of game attitude, at $p < 0.05$. However, there was no potential problem using the data in the analyses.

Results and discussion

Table 9 presents means and standard deviations of the three groups’ performances on the arithmetic test, GameBoy game, and attitude toward mathematics and the game. The students in Class

Table 7: Attitudes toward the technology game survey

Question	YES	NO
1. I know how to play the math game very well	YES	NO
2. I like the color and sound of the math game.	YES	NO
3. I like to know my scores of the game.	YES	NO
4. The scores help me to understand how well I did in the game.	YES	NO
5. I can control the math game options (eg. speed, level of difficulty)	YES	NO
6. I wish we would do more math game.	YES	NO
My most favorite thing about the math game is		
My least favorite thing about the math game is		

Table 8: Variable correlations ($n = 50$)

	Game scores	Game attitude	Math attitude	Gender
Game scores ($n = 49$)	—			
Game attitude ($n = 47$)	0.13	—		
Math attitude ($n = 48$)	0.16	0.36*	—	
Gender ($n = 50$)	-0.25	-0.07	-0.03	—
Ethnicity ($n = 50$)	-0.17	-0.24	-0.11	-0.07

Note. Number of students in each variable is different because of missing data.

*Correlation is significant at the 0.05 level (two-tailed). Gender: 1 = male, 2 = female.

Ethnicity: 1 = White, 2 = African-American, 3 = Asian, 4 = Hispanic, 5 = Other.

Table 9: Mean scores and standard deviations of variables ($n = 50$)

	Class One ($n = 19$)		Class Two ($n = 21$)		Class Three ($n = 9$)	
	Mean (%)	SD	Mean (%)	SD	Mean (%)	SD
Test total (70)	45.16 (65)	15.13	48.43 (68)	16.92	46.78 (67)	15.86
Basic (50)	39.63 (80)	11.69	39.76 (79)	11.38	41.33 (83)	12.10
Advanced (20)	5.53 (27)	4.96	8.67 (42)	6.14	5.44 (27)	5.05
Gender						
Male	44.18 (64)	16.52	55.08 (78)	11.37	55.00 (79)	6.06
Female	46.50 (66)	13.97	37.63 (54)	19.52	40.20 (57)	18.81
Ethnicity						
White	46.69 (67)	14.37	48.07 (69)	17.20	60.33 (86)	1.15
African-American	24.00 (34)	—	19.00 (27)	—	34.25 (49)	15.92
Asian	57.67 (82)	1.53	60.75 (88)	4.19	55.00 (79)	—
Hispanic	—	—	—	—	48.00 (69)	—
Other	27.00 (39)	5.66	41.00 (59)	15.56	—	—
Game attitude (6)	5.59 (93)	0.51	5.81 (97)	0.40	5.67 (95)	0.46
Mathematics attitude (23)	12.89 (58)	5.72	20.52 (89)	2.82	18.11 (75)	4.70
Game score						
Total ¹	3143.74	1489.02	3811.67	2009.02	4910.89	2263.57
Correct ²	2579.05	1472.40	2505.10	1882.87	3286.33	1806.89

¹Total number of problems attempted.

²Number of problems correctly solved.

Two had slightly higher scores on the test than those of Class One and Class Three. Their mathematical and game attitude scores were higher than those of Class One and Class Three as well. Class One had the lowest test, game and mathematics attitude score. On the basis of those results, we assume that the implementation of the game in Class One might be problematic. Further investigation is necessary to explore why and how the Class One conditions for using the game were not effective compared with the other two conditions.

Female students ($n = 21$, $M = 41.62$, $SD = 17.00$) in all three groups performed significantly lower than male students ($n = 28$, $M = 50.79$, $SD = 13.85$) on the test: $t(47) = 2.08$, $p < 0.05$. Asian students had the highest scores among all ethnic groups and African-American students scored lowest. As shown in Table 3, the pretest scores for Class One and Class Two, in Study One, considered by gender and by ethnicity, revealed a similar pattern to those in Study Two (see Tables 3 and 9). We suspect that the performance differences between males and females, and among ethnic groups resulted from their prior experiences rather than from the gaming activity itself. That is, the differences seen at the pretest level for white, African-American and Asian students in

mathematics, and for boys and girls in mathematics, have been shown in other research; thus, students' use of a technology-based game did not "create" the differences in performance.

For the multiple regression analysis, the first step of the analysis used all the predictor variables simultaneously, including attitude toward mathematics, attitude toward the GameBoy game, game scores, gender and ethnicity in the model. The results show that two variables, game scores ($\beta = 0.43$, $t = 3.58$, $p = 0.001$) and attitude toward mathematics ($\beta = 0.37$, $t = 3.03$, $p = 0.004$), emerged as significant predictors of arithmetic test scores. To reduce the error associated with the inclusion of nonsignificant predictors in regression models, only the significant predictors from the first regression analysis were used in the second step of the analysis. In this step, game scores ($\beta = 0.48$, $t = 4.32$, $p = 0.000$) and attitude toward mathematics ($\beta = 0.40$, $t = 3.52$, $p = 0.001$) accounted for 45% ($p < 0.001$) of the variance in arithmetic test scores (see Table 10). The overall plot and normal probability plot of residuals indicated that the results of this study met the assumptions of multiple regression analysis. Additionally, the results of the cross-validation technique confirmed that the prediction equation of the arithmetic scores works for students other than those who were used to develop the equation.

Overall, the cumulative number of problems solved correctly, as indicated by the GameBoy game scores, and the students' attitude toward mathematics significantly correlated with the students' scores on the arithmetic final test. Attitude toward the GameBoy game, gender and ethnicity were not significant predictors of test scores. Students who had higher scores on the GameBoy game were more likely to achieve high scores on the arithmetic test regardless of gender, ethnicity or attitude toward the GameBoy game. Therefore, game performance did influence students' scores on the arithmetic test. However, because of the limited sample size, we did not run statistical analysis to compare student performance between the three groups. The results of no differences between the three groups on the final test scores require further research. The findings are consistent with those of Study One regarding the differences on the final test between CG and GB. We speculate that there was possibly a teacher effect, in which the teaching practices of the three teachers might have influenced student learning of arithmetic. Additional qualitative data (eg. classroom observation) is necessary to explore the relationship between the teaching practices of the teachers and student learning. A future study using randomized controlled trials with three groups and one teacher would provide information on whether the frequency and duration of using the game influences student learning independent of teacher effect.

Lastly, we analyzed students' responses on two open-ended interview questions (see Table 6). Students' responses when asked about their "favorite thing about the game" can be organized into three categories: motivation, learning and game features (see Table 11). Sixteen students

Table 10: First and second steps of simultaneous regression analysis for variables predicting arithmetic scores

Variable ($n = 50$)	B	Standard Error	Coefficient β	t	p
Game scores	0.01	0.00	0.43	3.58	0.001
Attitude toward math	1.10	0.36	0.37	3.03	0.004
Gender	-6.05	3.73	-0.19	-1.62	0.112
Attitude toward game	2.23	4.26	0.07	0.52	0.604
Ethnicity	0.02	1.49	0.00	0.01	0.991
Constant	12.85	25.06		0.51	0.611
First step regression analysis. $R^2 = 0.47$ ($p < 0.001$)					
Game scores	0.01	0.00	0.48	4.32	0.000
Math attitude	1.12	0.32	0.40	3.52	0.001
Constant	15.04	6.04		2.49	0.016
Second step regression analysis. $R^2 = 0.45$ ($p < 0.001$)					

Table 11: Interview results regarding most and least favorite aspects of GameBoy game

Question	Total (n = 49)	Class One (n = 19)*	Class Two (n = 21)	Class Three (n = 9)
My most favorite thing about the math game is				
Motivation				
Playing game	13	3	7	3
Having Fun	3	3		
Learning (eg. improving scores, moving up levels, getting 100% correct)	16	4	8	4
Game features				
Various game tasks in GameBoy	7	3	2	2
Making character	3	1	2	
Learner control	3	1	2	
Do not know and no response	4	4		
My least favorite thing about the math game is				
Motivation				
Too difficult	13	5	5	3
Losing games (eg. incorrectly answering questions)	5		5	
Cannot play more	2		1	1
Game features				
Too easy (eg. some level is too slow)	1			1
Creating a character	1		1	
Repeat	1	1		
Program errors (eg. software errors)	3	1	2	
Nothing	23	12	7	4

*Missing one student.

(33%) responded that they had fun playing the game. Sixteen students (33%) responded that they like the game because it helps their learning. Seventeen students (35%) reported that their favorite thing about the game was its features including various game tasks, creating their own character and learner control. Student responses suggest that the technology-based game promoted positive attitudes toward learning and motivation because it provided various options that students could choose based on their individual needs. To the question about their least favorite thing about the game, 20 students (41%) responded positively in the perspective of motivation. Eighteen students (37%) responded that the game tasks are difficult, so they could not get a perfect score. Two students (4%) reported that they did not have sufficient time playing the game. Responses suggest that the game tasks were designed at an appropriate level of challenge so that the students repeated the tasks until they moved up to the next level rather than gave up on the tasks. Six students (12%) had negative responses to some game features and program errors. They reported that the game was boring because they had to repeat the easy game before going on to the next level. Students who had a negative reaction to making their own character did not elaborate on the reasons they disliked the feature. The student feedback provided us with information for revising the game in two ways. First, when students need to repeat the same level task, the repeated task should focus only on items that the student answered incorrectly rather than repeating all items at that level. Second, the game needs to have more character options so that students have a wider variety from which to choose.

Conclusion and future directions

The purpose of this investigation was to explore the effects of game technology on student learning in mathematics. The results of two quantitative studies provide evidence that game

technology positively impacts elementary students' learning of arithmetic regardless of ability level. Study One showed that given two types of arithmetic games for practice, students who played a technology-based game outperformed those who played a paper-based game, and students who played more outperformed those who played the game less frequently, regardless of student ability level. Study Two provided evidence that the technology-based game influenced student's arithmetic learning. The two studies support claims made in previous literature that game technology incorporating essential game features could be an effective learning tool for students to manage the acquisition of new information, depending on their prior knowledge and experience, learning progress, learning style, preferences and needs (Condie & Munro, 2007; Cordova & Lepper, 1996; McFarlane *et al.*, 2002; Rosas *et al.*, 2002).

Based on these findings, we suggest three important areas for further study of the impact of game technology on student learning. First, further research is needed to examine how game features should be designed to promote individualized instruction for enhancing learning outcomes of complex cognitive tasks, student interest and engagement in gaming environments. In the current study, we investigated the effects of technology-based games on student learning but did not examine relationships between various game features and the learning process. For example, research should examine the extent to which task-specific goals influence engagement in game activity and enhance student interest. In addition, qualitative research is necessary to understand how each game feature impacts cognitive engagement. In *Skills Arena*, students created their own characters and names for playing the game. Future research needs to investigate how the character creation feature influences student engagement, motivation, interest and learning outcomes. Such research is critical to providing more specific guidelines for game design.

Second, identifying significant classroom variables is an important step in successfully implementing games in educational settings. Classroom environments are complex; they are influenced by teacher variables (eg. teaching style, experience in incorporating technology in instruction, attitude toward the use of games), student variables (eg. gender, ethnicity, learning styles, strategies, and prior experience with content, game playing, and technology) and classroom variables (eg. number of students, number of technology devices, type of technology). Further research is needed to understand how such variables influence learners' engagement in games and which circumstances are most advantageous for successfully implementing technology-based games in classrooms.

Third, researchers need to explore the effectiveness of technology-based games on higher order thinking skills such as creativity, problem solving and decision making. We designed the game used in this study to support drill and practice in arithmetic because even though it is a low-level skill, mastery is a critical learning goal at this grade level, and class time is typically devoted to engaging in drill and practice exercises in a paper format. More research is needed to provide guidance on how designers can develop individualized technology game environments for development of advanced skills. To do this, designers need to further develop games with higher order thinking as the target learning goal and fully utilizing the affordances of game technology vis-à-vis various learning theories. Lastly, more extensive study with a larger sample size and sampling procedure are necessary in order to generalize the results of this study about the effects of game technology on student learning.

Acknowledgements

The authors wish to thank the teachers who gave us valuable feedback and continuous support to collect high-quality data from their classes. The authors also wish to acknowledge the helpful comments of Joseph Krajcik on this paper.

Note

Some initial research findings in this paper were first reported in Shin, Norris & Soloway (2006).

References

- Becker, K. (2007). Digital game-based learning once removed: teaching teachers. *British Journal of Educational Technology*, 38, 3, 478–488.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: the classification of educational goal: handbook I. Cognitive domain*. New York: Longmans, Green.
- Blumenfeld, P., Kempler, T. & Krajcik, J. (2006). Motivation and cognitive engagement in learning. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 475–488). Cambridge: Cambridge University.
- Bransford, J. D., Brown, A. L. & Cocking, R. R. (Eds) (2000). *How people learn: brain, mind, experience, and school* (Expanded ed.). Washington, D.C: National Academy Press.
- Carter, J. F. (1993). Self-management: education's ultimate goal. *Teaching Exceptional Children*, 25, 3, 28–32.
- Chandra, V. & Lloyd, M. (2008). The methodological nettle: ICT and student achievement. *British Journal of Educational Technology*, 39, 6, 1087–1098.
- Christmann, E. P. & Badgett, J. L. (2003). A meta-analytic comparison of the effects of computer-assisted instruction on elementary students' academic achievement. *Information Technology in Childhood Education Annual*, 1, 91–104.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Condie, R. & Munro, R. (2007). *The impact of ICT in schools—a landscape review: Coventry, Becta*. Retrieved April 18, 2011, from <http://webarchive.nationalarchives.gov.uk/20101102103654/publications.becta.org.uk//display.cfm?resID=28221>
- Cordova, D. I. & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, 88, 715–730.
- Dempsey, J., Haynes, L., Lucassen, B. & Casey, M. (2002). Forty simple computer games and what they could mean to educators. *Simulation & Gaming*, 33, 2, 157–168.
- Duschl, R. A., Schweingruber, H. A. & Shouse, A. W. (Eds) (2007). *Taking science to school: learning and teaching science in grades K-8*. Washington, DC: National Academies Press.
- Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., et al. (2007) *Effectiveness of reading and mathematics software products: findings from the first student cohort* (Publication No. 2007-4005). Retrieved October 26, 2010, from Institute of Education Sciences, U.S. Department of Education, website: <http://ies.ed.gov/ncee/pdf/20074005.pdf>
- Eck, R. V. (2006). The effect of contextual pedagogical advisement and competition on middle-school students' attitude toward mathematics and mathematics instruction using a computer-based simulation game. *Journal of Computers in Mathematics and Science Teaching*, 25, 2, 165–195.
- Finn, M. & Vandenham, N. (2004). The handheld classroom: educational implications of mobile computing. *Australian Journal of Emerging Technologies and Society*, 2, 1, 1–15.
- Garris, R., Ahlers, R. & Driskell, J. E. (2002). Games, motivation, and learning: a research and practice model. *Simulation & Gaming*, 33, 4, 441–467.
- Hoon, T. S., Chong, T. S. & Binti Ngah, N. A. (2010). Effect of an interactive courseware in the learning matrices. *Educational Technology & Society*, 13, 1, 121–132.
- Inal, Y. & Cagiltay, K. (2007). Flow experiences of children in an interactive social game environment. *British Journal of Educational Technology*, 38, 3, 455–464.
- Kahveci, M. & Imamoglu, Y. (2007). Interactive learning in mathematics education: review of recent literature. *Journal of Computers in Mathematics and Science Teaching*, 26, 2, 137–153.
- Kernan, M. C. & Lord, R. G. (1990). Effects of valence, expectancies, and goal-performance discrepancies in single and multiple goal environments. *Journal of Applied Psychology*, 75, 194–203.
- Kinzie, M. B. & Joseph, D. R. D. (2008). Gender differences in game activity preferences of middle school children: implications for educational game design. *Education Technology Research & Development*, 56, 5/6, 6443–6663.
- Locke, E. A. & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.
- McFarlane, A., Sparowhawk, A. & Heald, Y. (2002). *Report on the educational use of games: an exploration by TEEM of the contribution which games can make to the education process*. Retrieved April 18, 2011, from www.teem.org.uk/publications/teem_gamesined_full.pdf

- Means, B., Toyama, Y., Murphy, R., Bakia, M. & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: a meta-analysis and review of online-learning studies*. Washington, DC: U.S. Department of Education. Retrieved October 27, 2010, from <http://www.educause.edu/Resources/EvaluationofEvidenceBasedPract/174235>
- Miller, D. J. & Robertson, D. P. (2010). Using a games console in the primary classroom: effects of "Brain Training" programme on computation and self-esteem. *British Journal of Educational Technology*, 41, 2, 242–255.
- Morrison, G. R., Ross, S. M. & Baldwin, W. (1992). Learner control of context and instructional support in learning elementary school mathematics. *Educational Technology, Research, & Development*, 40, 5–13.
- Norris, C., Shin, N. & Soloway, E. (2007). Educational technology for the mainstream: a call for designing for simplicity and reliability. *Educational Technology Magazine*, 47, 3, 6–9.
- Pearce, K., Lungren, M. & Wince, A. (1999). The effects of curriculum practices on first graders' attitudes, activity preference, and achievements in mathematics. *Education*, 119, 1, 82–90.
- Piaget, J. & Inhelder, B. (1967). *The Child's Conceptions of Space*. Compton Printing Ltd.
- Quintana, C., Shin, N., Norris, C. & Soloway, E. (2006). Learner-centered design: reflections on the past and directions for the future. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 119–134). Cambridge: Cambridge University.
- Randel, J., Morris, B., Wetzel, C. & Whitehill, B. (1992). The effectiveness of games for educational purposes: a review of recent research. *Simulation & Gaming*, 23, 261–276.
- Ricci, K., Salas, E. & Cannon-Bowers, J. A. (1996). Do computer-based games facilitate knowledge acquisition and retention? *Military Psychology*, 8, 4, 295–307.
- Rieber, L. P. (1996). Seriously considering play: designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development*, 44, 43–58.
- Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P., et al. (2002). Beyond Nintendo: design and assessment of educational video games for first and second grade students. *Computers & Education*, 40, 71–94.
- Rose, D. H., Meyer, A. & Hitchcock, C. (2005). *The universally designed classroom: accessible curriculum and digital technologies*. Cambridge, MA: Harvard Education Press.
- Shin, N., Norris, C. & Soloway, E. (2006). Findings from early research on one-to-one handheld use in K-12. In M. van't Hooft & K. Swan (Eds), *Ubiquitous computing in education* (pp. 19–39). Mahwah, NJ: Lawrence Erlbaum Associates.
- Shin, N., Norris, C., & Soloway, E. (2006, June). *Effects of handheld games on students learning*. Paper presented at the annual meeting of the International Conference of Learning Sciences, Bloomington, IN.
- Shin, N., Stevens, S. Y. & Krajcik, J. (2010). Tracking student learning over time using Construct-Centered Design. In S. Routledge (Ed.), *Using analytical frameworks for classroom research: collecting data and analysing narrative* (pp. 38–68). London: Taylor & Francis.
- Sutherland, L. M., Shin, N. & McCall, K. L. (2010). *Integrating science inquiry, literacy, technology, and universal design for learning to enhance middle school students' opportunities to learn science*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Philadelphia, PA.
- Thomas, P. & Macredie, R. (1994). Games and the design of human-computer interfaces. *Educational and Training Technology International*, 31, 2, 134–142.
- Urdu, T. C. & Maehr, M. L. (1995). Beyond a two-goal theory of motivation and achievement: a case for social goals. *Review of Educational Research*, 65, 3, 213–243.
- Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K. & Wright, M. (2006). Computer gaming and interactive simulations for learning: a meta-analysis. *Journal of Educational Computing Research*, 34, 3, 229–243.
- Vonesh, E. F. & Chinchilli, V. G. (1997). *Linear and nonlinear models for the analysis of repeated measurements*. London: Chapman and Hall.
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wagner, E. D. (1994). In support of a functional definition of interaction. *American Journal of Distance Education*, 8, 2, 6–29.
- Whitehall, B. & McDonald, B. (1993). Improving learning persistence of military personnel by enhancing motivation in a technical training program. *Simulation & Gaming*, 24, 294–313.
- Wolters, C. A., Yu, S. L. & Pintrich, P. R. (1996). The relation between goal orientation and students' motivational beliefs and self-regulated learning. *Learning and Individual Differences*, 8, 3, 211–238.