A Computer-Based Game that Promotes Mathematics Learning More than a Conventional Approach

Bruce M. McLaren, Carnegie Mellon University, Pittsburgh, PA, USA

Deanne M. Adams, Notre Dame, USA

Richard E. Mayer, University of California, Santa Barbara, CA, USA

Jodi Forlizzi, Carnegie Mellon University, Pittsburgh, PA, USA

ABSTRACT

Excitement about learning from computer-based games has been palpable in recent years and has led to the development of many educational games. However, there are relatively few sound empirical studies in the scientific literature that have shown the benefits of learning mathematics from games as opposed to more traditional approaches. The empirical study reported in this paper provides evidence that a mathematics educational game can provide superior learning opportunities, as well as be more engaging. In a study involving 153 students from two middle schools, 70 students learned about decimals from playing an educational game—Decimal Point—whereas 83 students learned the same content by a more conventional, computer-based approach. The game led to significantly better gain scores in solving decimal problems, on both an immediate ($d = .43$) and delayed ($d = .37$) posttest and was rated as significantly more enjoyable ($d = .95$). Low prior knowledge students especially benefitted from the game. This paper also summarizes the game's design characteristics.

KEYWORDS

Computer Games, Digital Games, Educational Games, Mathematics Learning, Mathematics Problem Solving

INTRODUCTION

The enthusiasm about computer-based educational games is by now well documented and widespread. Many claims have been made about the benefits of learning with educational games versus more traditional approaches (Gee, 2003; Prensky, 2006; Squire & Jenkins, 2003). Furthermore, teachers believe that computer-based games can be effective. For instance, a 2014 survey found that 55% of 513 teachers who use games in the classroom use them at least once a week (Gamesandlearning. org, 2015). Given the obvious appeal of computer-based games more generally – the computer game industry is growing much faster than the U.S. economy as a whole (Siwek, 2010) and 97% of students aged 12 through 17 play video games regularly (Lenhart et al, 2008) – it is easy to understand and embrace the enthusiasm about and promise of computer games as a way to engage kids and lead to meaningful learning.

Yet, while strong claims have been made about the potential of educational computer games, those claims are, thus far, based on relatively weak evidence (Hannifin & Vermillion, 2008; Honey & Hilton, 2011; Mayer, 2014; O'Neil & Perez, 2008; Tobias & Fletcher, 2011). For instance, Mayer (2014) extensively collected and evaluated the published scientific evidence in which an educational game was compared to a more traditional instructional approach (so-called media comparison studies).

He eliminated all of the studies that did not meet rigorous scientific study criteria, such as comparing an experimental (game) and control (non-game) condition with the same academic content, inclusion of a dependent measure that involves academic outcome, and reports of means, standard deviations, and sample sizes for the learning outcomes. Mayer's evaluation uncovered only 16 rigorous studies in science and 5 in mathematics[1]. While 12 of the 16 studies in science showed learning benefits for the games group (mean $d = 0.69$), only 3 of the 5 studies in math showed learning benefits for the games, with a negligible effect size of 0.03.

Other meta-analyses of educational games have reported positive results for educational games more generally, but not for mathematics educational games more specifically (Clark, Tanner-Smith, & Killingsworth, 2015; Sitzmann, 2011; Vogel et al., 2006; Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013). For instance, Clark et al. (2015), in a review of 69 sound, empirical studies (filtered from over 1000 studies reported in published papers), found that computer-based educational games were associated with a 0.33 standard deviation improvement over non-game comparison conditions. Clark et al. (2015) emphasize that educational games are designed in many different ways, vary on a variety of dimensions, so they argue more for the importance of how the variations in game designs lead to different learning outcomes (called value-added studies of games by Mayer, 2014) and less on media comparisons within content domains (e.g., mathematics). Thus, they do not separately evaluate the evidence of digital games in the domain of mathematics. However, they reach the same general conclusion of Mayer, saying: "methodological rigor needs to be increased in research on games for learning" (Clark et al., 2015, pp. 35).

In other words, the research field of educational games is still nascent, with limited empirical evidence about the effectiveness of games, especially in the domain of mathematics. In fact, educational technology researchers have only recently begun to investigate ways to inject the learning of traditional subjects into computer games (Aleven, Myers, Easterday, & Ogan, 2010; Conati & Manske, 2009; Habgood & Ainsworth, 2011; Lomas, Patel, Forlizzi, & Koedinger, 2013; McNamara, Jackson, & Graesser, 2010; Risconscente, 2013). While there is certainly reason for the excitement about educational games, and many educational technologists are increasingly investigating the potential of games for learning, it is important that more rigorous studies be conducted to determine whether the excitement is justified. The educational game and study presented in this paper represents a step in that direction.

DECIMAL POINT: A GAME FOR LEARNING DECIMALS

Our educational game, a single-player game that has been under development for three years (Forlizzi et al., 2014), is based on an amusement park metaphor and is targeted at middle-school students. The game is called "Decimal Point: The Fantastically Fabulous World of Fractional Fun" (for short, "Decimal Point"). As shown in Figure 1, the student travels sequentially to different theme areas (e.g., Haunted House, Wild West, Space Adventure, Old-time Amusement Park), playing a variety of mini-games within each theme area targeted at learning decimals and relevant to that area's leitmotif (e.g., Enter If You Dare in the Haunted House; Bronco Lasso in the Wild West; Space Raider in Space Adventure; Balloon Pop! in the Old-time Amusement Park). The student's progress through the park is tracked, and students are visually prompted for the next game they will play. In Figure 1, the student has already played all of the mini-games up to but not including the Old-Time Amusement Park, indicated by the colored circles. The student is prompted to pick the next mini-game, Balloon Pop!, indicated by the pulsating red circle around that mini-game (see the middle of Figure 1).

Figure 1. The Game - Decimal Point: The Fantastically Fabulous World of Fractional Fun
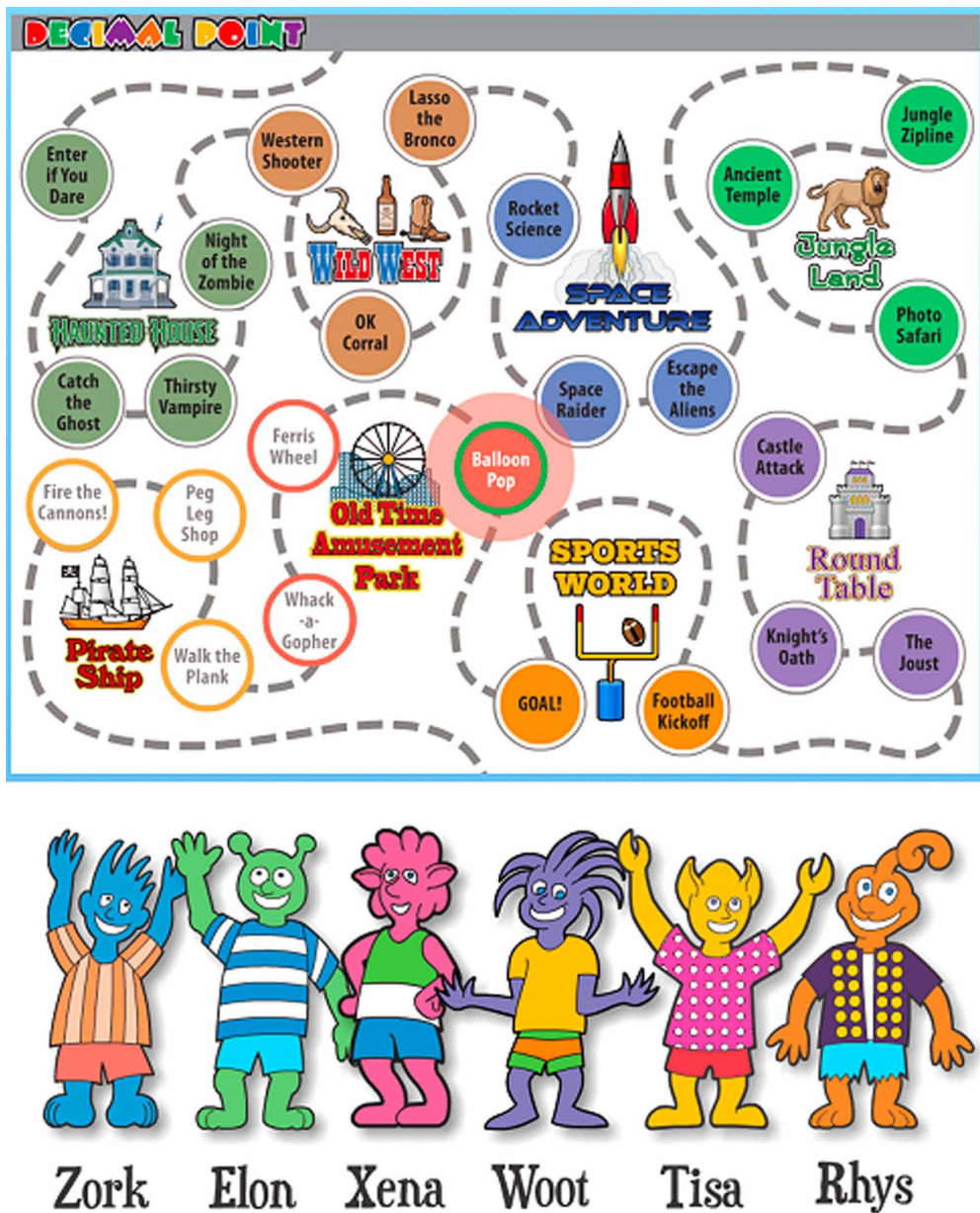
At the outset of game play, the students are introduced to a group of six fantasy characters – Zork, Elon, Xena, Woot, Tisa, and Rhys, shown at the bottom of Figure 1. The fantasy characters act as guides to Decimal Point and encourage students to play and congratulate them when they correctly solve problems. Students do not score points or compete with their fellow students. Rather, students simply make their way through the entire amusement park and are commended upon completing the journey by the fantasy characters (see Figure 2).

Figure 2. The end of Decimal Point

The problems within the mini-games are focused on common decimal misconceptions (Glasgow, Ragan, Fields, Reys, & Wasman, 2000; Isotani, McLaren, & Altman, 2010; Irwin, 2001; Stacey, Helme, & Steinle, 2001). Decimals are critical for students to master in order to advance to more complex mathematics (National Mathematics Advisory Panel, 2008) and many people struggle with decimals into adulthood (Putt, 1995). An example Decimal Point mini-game is shown in Figure 3. The "Balloon Pop!" mini-game challenges the student to toss darts at decimal-labeled balloons (e.g., 0.49, 0.2, 0.1921, 0.382) in the order from smallest to largest decimal. This game is targeted at the common misconception in which students think longer decimals are larger than shorter decimals (e.g., 0.1921 > 0.2), presumably due to their prior experience with whole numbers (Stacey et al., 2001). (A more extensive discussion of decimal misconceptions and the ways those misconceptions are tested within the instructional materials are provided in the next section.) The student tries to hit the balloons in the requested order and, if they make mistakes, is prompted to correct their solution by dragging and dropping the numbers into a new sequence. Besides ordering decimals, the various mini-games challenge students with other types of decimal problems, including to correctly placing a point on a number line, completing decimal sequences, and placing decimals in less-than and greater-than "buckets" in comparison to a given decimal.

After playing a mini-game and correctly solving the problem, the student is prompted to explain his or her solution, by choosing possible explanations from a multiple-choice list. This is done in keeping with research that has demonstrated the learning benefits of prompted self-explanation (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, DeLeeuw, Chiu, & LaVancher, 1994), including prompted self-explanation in games (Johnson & Mayer, 2010; Mayer & Johnson, 2010). Figure 4

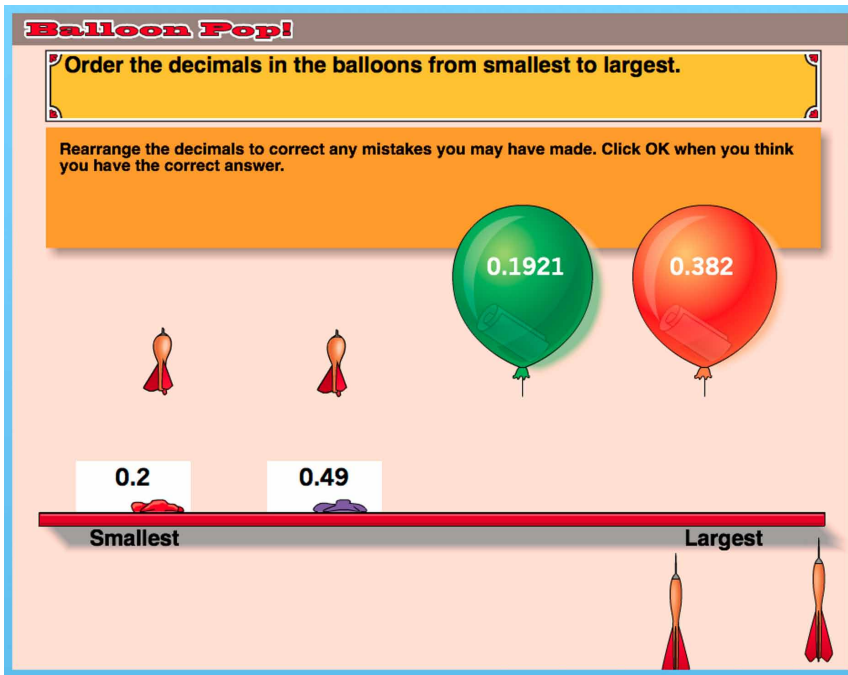**Figure 3. A student playing the "Balloon Pop!" mini-game**

Figure 4. A student explaining their answer in the "Enter If You Dare" mini-game

shows a student prompted to explain their solution in another mini-game, "Enter If You Dare", where the student has been prompted to place a skull at the correct point on a number line in order to enter a haunted house.

The non-game version of the instructional materials presents a conventional user interface to the students, prompting them to solve decimal problems. In the non-game version of the decimal ordering problem, the student is presented with a list of decimals and prompted to drag and drop the decimals to the correct order (which has already been done in Figure 5). The student is then prompted to explain his or her solution (see the bottom of Figure 5), just as the player of the "Balloon Pop!" mini-game is prompted for explanation after playing the mini-game and correctly ordering the decimals.

For the media comparison study described in this paper, we presented students in the game condition with mini-game problems that are identical in content to the corresponding conventional items in the non-game condition. For example, the mini-game of Figure 3 and the conventional problem of Figure 5 occur at the same point in the instructional sequence across the two conditions. In both items, the student is prompted to order the four decimals (0.49, 0.2, 0.1921, 0.382) and explain how they ordered the decimals. However, in the mini-game the student is presented with a short narrative and a fictional context in which the game is situated. The problem-solving activity is embedded playfully in the game context (e.g., in Figure 3, the player pops the balloons by "throwing" a virtual dart to reveal the decimal values to order). In the non-game context, the problems are solved in a more straightforward and conventional manner, with a minimalist user interface. This approach is consistent with Clark's (2001) admonishment for media comparison studies: to make sure that the instructional content (e.g., the specific decimals) and instructional methods (e.g., solving problems and giving explanations) are equivalent.

Figure 5. A student solving a conventional problem-solving item that is equivalent to the "Balloon Pop!" mini-game

All of the instructional materials, both game and non-game, run on the web, within a standard browser, and were developed using Flash and the Cognitive Tutor Authoring Tools (CTAT; Aleven, McLaren, Sewall, & Koedinger, 2009). The materials have been deployed on a web-based learning management system called TutorShop, which presents the materials in a given sequence and logs all student actions. Forty-eight decimal problems (two problems per mini-game and conventional interface) have been implemented for the game and non-game curricula.

THE IMPORTANCE OF DECIMALS IN MATHEMATICS

Our materials, both within the educational mini-games and conventional problem-solving environment, are aimed at helping middle school students learn decimals – a fundamental area of mathematics with common misconceptions (Glasgow et al., 2000; Irwin, 2001; Stacey et al., 2001) that often persist into adulthood (Putt, 1995). Decimals, and more generally rational numbers, are identified as a fundamental gateway topic in mathematics (Kouba, Carpenter, & Swafford, 1989; National Mathematics Advisory Panel, 2008; Rittle-Johnson, Siegler, & Alibali, 2001). Decimal misconceptions have been shown to lead to difficulty in learning more advanced mathematics (Hiebert & Wearne, 1985). Also, the Common Core State Standards Initiative (http://www.corestandards.org/) lists decimals as an important topic, essential for advancing to more complex mathematical subjects.

Through an extensive review of the math education literature, we have compiled a taxonomy of 17 common and persistent decimal misconceptions (Isotani et al., 2010), covering 32 published papers and extending as far back as 1928 (e.g., Brueckner, 1928; Glasgow et al., 2000; Graeber & Tirosh, 1988; Irwin, 2001; Resnick et al., 1989; Sackur-Grisvard & Léonard, 1985; Stacey et al., 2001). Most of the 32 published papers address either a single misconception or a small set of related misconceptions.

Table 1. Decimal misconceptions that are targeted with the instructional materials

Brief description	Decimal tests to probe for the misconception	References
Longer decimals are larger	"Place 0.34 on a number line that already has 0.1, 0.3, and 0.4 on it" (placement to the right of 0.4 provides evidence for "Longer decimals are larger")	Hiebert et al., 1991; Irwin, 2001; Resnick et al., 1989; Rittle-Johnson et al., 2001; Sackur-Grisvard & Léonard, 1985; Stacey et al., 2001
	"Choose the largest of the following numbers: 0.6754, 0.78, 0.8, 0.321" (choosing any answer but "0.8" – but especially 0.6754 – is evidence for "Longer decimals are larger")	
Shorter decimals are larger	"Order the following decimals, smallest to largest: 0.721, 0.3, 0.42" (ordering of 0.721, 0.42, 0.3 would indicate "Shorter decimals are larger")	Irwin, 2001; Putt, 1995; Sackur-Grisvard & Léonard, 1985; Stacey et al., 2001
	"Select the list that goes from largest to smallest" 0.1 0.89 0.333 0.2214 (→ "Shorter decimals are larger" misconception) 0.89 0.333 0.2214 0.1 (→ Correct) 0.2214 0.333 0.89 0.1 (→ "Longer decimals are larger" misconception)	
The numbers to the left and the right of the decimal point are separate and independent	"What is the answer to 34.53 + 3.5?" (A student with this misconception may not carry, thinking the decimal acts as an 'invisible boundary' between the left and right side of the decimal and thus could come up with an incorrect result such as 37.103 or 37.58.)	Irwin, 2001
	"Complete the following sequence: 0.3, 0.6, 0.9, ___, ____." (A student who fills in "0.12" and/or "0.15" may have this misconception.)	
Decimals less than 1.0 are less than zero	"Order the following decimals, smallest to largest: 1.3, 0.1, 0" - Ordering of 0.1, 0, 1.3 would provide evidence for "Decimals less than 1.0 are less than zero"	Irwin, 2001; Putt, 1995; Stacey et al., 2001
	"If a decimal number starts with a 0 before the decimal point, is it < 0? o Yes o No o It Depends o Don't Know" (An answer of "Yes" is evidence for "Decimals less than 1.0 are less than zero".)	

The four most common misconceptions found in the literature are shown in Table 1. The most common misconception is the "longer decimals are larger" misconception illustrated in Figures 3 and 5 and which is likely the result of students learning whole numbers before learning decimals. The decimal instructional materials that we have developed and used with over 1,000 students in a variety of studies over the past five years (cf. Adams et al., 2014; McLaren, Adams, & Mayer, 2015) are explicitly targeted at these four misconceptions. These four misconceptions are also the target of the instructional materials of the present study.

RESEARCH QUESTIONS AND HYPOTHESES

We had the following research questions in conducting this media comparison study.

RQ1: *Does the game lead to better decimal learning than the more conventional (i.e., non-game) instructional approach?* The most basic question we had was whether Decimal Point would lead to superior learning benefits compared to the conventional instructional materials. Our hypothesis was that the game, while probably more enjoyable for the students, would not lead to better learning gains. This hypothesis was supported by the Mayer (2014) findings showing that math games generally have not led to better learning, as well as the straightforward notion that both groups would receive the same instructional content and instructional method, i.e., practice on solving and explaining the same decimal problems. On the other hand, if the motivating properties of the game encourage students to exert more effort in learning, then we would expect the game group to outperform the non-game group in decimal problem solving.

RQ2: *Do students in the game group enjoy their learning experience more than the students in the non-game group?* While we did not expect better learning gains for the game, we hypothesized that students would enjoy learning with the game more than with the non-game, because of the more playful, colorful, and alluring details of the game. In addition, various prior research has demonstrated that students experience heightened engagement and motivation when they play educational games (Ke, 2008; Schaaf, 2012; Tarng & Tsai, 2010).

RQ3: *Do the students in the game group make more errors than the students in the non-game group?* While interested in this question, we did not have a directional hypothesis with respect to the errors the students might make in completing the game and in solving the non-game problems. On one hand, the allure of the game might have improved students' attention and focus, thus leading to fewer errors. On the other hand, the game might have provided a distraction that led to less concentration on the mathematics content and instruction and thus produced more errors.

RQ4: *Does the game take longer to complete than the non-game?* Because of the alluring details and additional steps in the game (e.g., a brief, motivational introduction to the game; the amusement park map shown between playing mini-games (see Figure 1); the less direct way of solving the problems in the mini-games; the brief congratulatory video at the end (see Figure 2)), we expected the game would take students longer to complete than the non-game. We examined this question as a practical issue because the potential benefits of game playing must be weighed against the cost of additional time required. Put another way, if students learn more with the game, but it takes them much longer, an "opportunity cost" of playing the game is paid.

RQ5: *Do students in the game group gain more confidence in the domain (i.e., decimals) than the game group after they complete the materials?* We hypothesized that the game would give students more confidence in their decimal knowledge and ability after they made their way through the amusement park, due to the fun, motivation, and positive disposition the game provides.

METHOD

Participants and Design

The original set of participants were 213 6th grade students from two middle schools (one urban, one suburban) in the U.S. Fifty-two (52) participants (26 game, 26 non-game) were excluded from analyses because they did not fully complete all materials and measures in the study. In addition, 8 participants were removed due to having gain scores 2.5 standard deviations above or below the mean between the pretest and the immediate posttest or between the pretest and the delayed posttest. The remaining 153 students (66 male, 87 female) had a mean age of 11.3 (SD = .52). Because of the distraction and demotivation that might have occurred with students sitting next to one another working with different materials, participants were assigned by classroom to one of the two instructional conditions: (1) Game (n = 70) or (2) Non-Game (n = 83). The eleven classes across the two schools were not randomly assigned to the Game and Non-Game groups. Instead, they were assigned to balance low performing and high performing classes between the Game and Non-Game groups.

Materials

A web-based learning environment was used to deploy the experiment, and instructional materials were assigned to each group as outlined in Table 2. Materials included three tests (pretest, posttest, delayed posttest), the game or non-game materials, and two questionnaires (demographic, evaluation). Details about the materials are provided in the remainder of this section.

Pretest, Posttest, and Delayed Posttest. The pretest, posttest, and delayed posttest, was administered online and consisted of twenty-four items. Some of the twenty-four items had multiple parts, comprising 61 possible points. Participants received points for each correct part. There was an A, B, and C form of the base test, which were isomorphic to one another and which were positionally counter-balanced within condition (i.e., approximately 1/3 of the students in each condition received Test A as the pretest, 1/3 received Test B as the pretest, and 1/3 received Test C as the pretest; likewise for the posttest and delayed posttest).

Test items were designed to probe for specific decimal misconceptions and took a variety of forms, including adding, multiplying, and dividing decimal numbers (e.g., $0.387 + 0.05 = \_\_\_\_$, $6.5 \times 100 = \_\_\_\_$), choosing the largest of a given set of decimals (e.g., "Choose the largest of the following three numbers: 5.413, 5.75, 5.6"), placing a given decimal number on a number line, answering conceptual

Table 2. Conditions and Materials used in the study. Italicized items vary across conditions

Game (Experimental Condition)	Non-Game (Control Condition)
Pretest (A, B, or C)	Pretest (A, B, or C)
Demographic Questionnaire	Demographic Questionnaire
Intervention - Game	Intervention – Non-Game
Game-Item-1a	Non-Game-Item-1a
Game-Item-1b	Non-Game-Item-1b
…	…
Game-Item-24a	Non-Game-Item-24a
Game-Item-24b	Non-Game-Item-24b
Evaluation Questionnaire	Evaluation Questionnaire
Posttest (A, B, or C)	Posttest (A, B, or C)
Delayed Posttest (A, B, or C)	Delayed Posttest (A, B, or C)

Figure 6. Example three-part test item

questions (e.g., "Is 786 / 987 less than zero, zero, or greater than zero?", "Convert the following verbal decimal to numeric form: 32 hundredths"), and completing decimal sequences (e.g., "Write down the next item in the following sequence: 0.201, 0.401, 0.601, 0.801", ____).

As mentioned, multiple points were scored for some items by, for instance, multiple problems being presented as sub-parts of a single item (i.e., on a single screen). For instance, the "Write down the next item …" problem was displayed with two other similar decimal sequence problems on the same screen (See Figure 6). Students could advance without answering items by selecting the button in the lower right at any time. After a student had answered problems on a single screen and moved on to the next item on the test, by selecting "Finish" on the screen, they could not return to that item. Grading was done to recognize different possible correct answers (e.g., for the second problem of Figure 6, the answers "0.98", ".98", "0.980" would all be graded as correct.).

- **Demographic Questionnaire.** Immediately after the pretest, students were asked standard demographic questions (e.g., age and gender) and also questions related to their prior knowledge of and work with decimals, experience working with computers, and questions relating to math self-efficacy. The complete demographic questionnaire is shown in Figure 7.
- **Confidence-rating scale.** After eleven of the answers on the pretest, posttest, and delayed posttest, students were prompted to rate their confidence in the just-answered item using a 5-point Likert

scale, i.e., "How sure are you that you solved this problem correctly? Not at all, A little sure, Somewhat sure, Sure, Very Sure".

- **Evaluation Questionnaire.** Upon completing the intervention, students were given an online evaluation questionnaire to rate various aspects of their experience with the materials. The complete demographic questionnaire is shown in Figure 8. The questionnaire included 11 items that were later combined into four categories:
  ◦ *Lesson Enjoyment* (i.e., how well students liked the lesson). Two items: "I liked doing this lesson", "I would like to do more lessons like this";
  ◦ *Ease of Interface Use* (i.e., how easy it was for the student to interact with the intervention items and interface). Four items: "I liked the way the material was presented on the screen", "I liked the way the computer responded to my input", "I think the interface of the system was confusing", "It was easy to enter my answer into the system";
  ◦ *Feelings of Math Efficacy* (i.e., whether the student had positive feelings about mathematics after using these materials). Two items: "The lesson made me feel more like I am good at math", "The lesson made me feel that math is fun";
  ◦ *Perceived Material Difficulty* (i.e., whether the student perceived that the lesson was difficult). Three items: "The material in this lesson was difficult for me", "I worked hard on understanding the material in this lesson", "I could easily understand the assignment"[2]. As shown in Figure 8, responses were given using a 5-point Likert scale ranging from "Strongly agree" (1) to "Strongly disagree" (5).
- **Intervention Items:** A total of 48 intervention items (24 pairs, see Table 2) were presented online to both the Game and Non-Game participants. Each consecutive pair of items (e.g., see Table 2: Game-Item-1a and Game-Item-1b) involved the same type of mini-game or conventional item. However, a different specific decimal problem was presented in each item of the pair. For instance, the Balloon Pop! mini-game of Figure 3 and the equivalent conventional problem of Figure 5 were in the same item position in the Game and Non-Game intervention sequences, respectively. Thus, the Game participants played two iterations of each of the 24 mini-games of Figure 1 (yielding a total of 48 decimal problems), while the Non-Game participants solved 48 equivalent conventional problems. In both conditions correct responses turned green and incorrect responses turned red. No hints or error feedback messages were provided in either condition.

PROCEDURE

The experiment was conducted at the students' schools as replacement for their regular mathematics classes. In total, the study took seven lesson periods, each of approximately 45 minutes in length, to complete. Students received a login for the web-based environment and could work at their own pace through the materials. When they finished the first posttest, however, they could not progress to the delayed posttest; this test took place one week later, on the seventh and final period. The phases and tasks they encountered are summarized in Table 2.

RESULTS

We first analyzed whether the participants in the Game and Non-Game conditions were balanced in terms of gender, age, self-reported math proficiency, and pretest scores. The Game and Non-Game conditions did not differ significantly in terms of the distribution of males and females, χ^2 (1, $N = 153$) = .07, $p = .79$; age, t (150) = 1.57, p -.12. However, an ANOVA revealed that students in the Game condition scored significantly higher on the pretest than students in the Non-Game condition, F (1,151) = 12.85, $p < .001$. In addition, participants in the Game group tended to report higher familiarity with decimals compared to the Non-Game group, $t(151)$= -2.9, $p = .004$, and had higher self-reported math proficiency, $t(151)$= -2.5, $p = .01$. The overall results are presented in Table 3.

Figure 7. The demographic questionnaire

Figure 8. The evaluation questionnaire

Table 3. Mean values of dependent measures per condition

	Game (n=70)	*Non-Game (n=83)*	*Effect Size (d)*	*Effect Size (d) for adj. means*
Pretest (max=61)	32.0 (11.0) *	26.0 (9.7)	0.58	-
Posttest (max=61)	38.6 (9.6) * ^	30.1 (10.5)	0.84	0.65
Delayed posttest (max=61)	40.6 (9.3) * ^	31.9 (11.2)	0.84	0.59
Pretest-to-posttest gain scores	6.9 (6.9) * ^	4.1 (6.1)	0.43	0.68
Pretest-to-delayed posttest gain scores	8.6 (7.5) * ^	5.9 (7.1)	0.37	0.59
Lesson enjoyment (1-5)	2.0 (1.0) *	3.0 (1.1)	0.95	-
Ease of interface (1-5)	2.1 (.69) *	2.5 (.76)	0.55	-
Feelings of math efficacy (1-5)	2.1 (1.0) *	2.8 (1.0)	0.80	-
Intervention time (minutes)	89.7 (28.5)	82.1 (25.4)	0.28	-
Intervention errors	175.0 (111.4)	273.4 (209.2) *	0.57	-

Asterisks (*) indicate scores that are significantly greater for non-adjusted means. Carrot tops (^) indicate scores that are significantly greater for adjusted means (i.e., with the pretest score used as a covariate).

Because of the unbalanced pretests, we ran the test analyses with the pretest as a covariate and report those effect size results in the final column of Table 3. Results related to each of the research questions are presented and discussed below.

RQ1: *Does the game lead to more decimal learning than the more conventional (i.e., non-game) instructional approach?* Using MANCOVA to assess the posttests, with the pretest as covariate, the Game group performed significantly better than the Non-Game group on both the posttest ($F(1,150) = 15.32, p < .001, d = .65$ for adjusted means) and delayed posttest ($F(1,150) = 12.53, p < .001, d = .59$ for adjusted means). When looking at gain scores, shown in Table 3 and also using MANCOVA, the same significant results were found (i.e., Posttest: Game group over Non-Game group, $F(1,150) = 15.32, p < .001, d = .68$ for adjusted mean gains; Delayed posttest: Game group over Non-Game group, $F(1,150) = 12.53, p = .001, d = .59$ for adjusted mean gains). The superiority of the Game group on tests of learning outcome is the main finding in this study.

RQ2: *Do students in the game group enjoy their learning experience more than the students in the non-game group?* Using independent sample t-tests to assess three of the questionnaire evaluation categories (i.e., Lesson Enjoyment, Ease of Interface Use, Feelings of Math Efficacy) it was determined, as hypothesized, that the game group had significantly more positive feelings about their experience after completing the intervention. The Game group reported (a) liking the intervention significantly more $t(151)= -4.4, p < .001$ (b) thinking the intervention interface was easier to use $t(151)=-3.2, p = .002$, and (c) having more positive feelings about math $t(151) = -5.7, p < .001$.

RQ3: *Do the students in the game group make more errors during learning than the students in the non-game group?* The Non-Game group made significantly more errors during the intervention than the Game group, $t(151) = -3.5, p = .001$. We were interested in this question, but did not have a hypothesis about it.

RQ4: *Does the game take longer to complete than the non-game?* The Game group took longer, but not significantly so, to complete the intervention, $t(151)=1.7, p =.08$. Thus, our hypothesis was not confirmed that the Game group would take longer to complete the intervention than the Non-Game group.

RQ5: *Do students in the game group gain more confidence in the domain (i.e., decimals) than the game group as they complete the materials?* Using the self-reported confidence ratings after 11 of the items on each of the tests and based on repeated measures ANOVAs, all students got more confident over the course of the three test sessions ($F(2,292) = 27.81$, $p < .001$). However, there was not a significant interaction for session x condition ($F(2, 292) = .68$, $p = .51$). Thus, the hypothesis that the Game participants would gain more confidence in decimals than the Non-Game participants was not confirmed.

Finally, while it was not one of our original research questions, we checked whether the interventions were more beneficial to high or low prior knowledge students. Students were classified as having high or low prior knowledge using a median split on the pretest. For the Game group, 43 of the students were classified as high while 27 were classified as having low prior knowledge. For the Non-Game condition, 34 students were classified as having high prior knowledge while 49 were classified as having low. A MANOVA revealed that prior knowledge had a significant effect on gains scores between the pretest and immediate posttest, $F(1,49) = 19.76$, $p < .001$, as well as on gain score between the pretest and delayed posttest $F(1,149) = 15.44$, $p < .001$. Students who were classified as having lower prior knowledge showed higher gain scores (Pre-Immediate Posttest $M = 7.05$ $SD = 7.62$; Pre-Delayed Posttest $M = 8.83$ $SD = 8.32$) compared to the high prior knowledge students (Pre-Immediate Posttest $M = 3.74$ $SD = 4.99$; Pre-Delayed Posttest $M = 5.49$ $SD = 5.82$), suggesting that both interventions were more effective for students who had difficulty with decimal problems. In addition, there was a significant interaction between the Game and Non-Game groups and prior knowledge level for both the gains scores between the pretest and immediate posttest, $F(1,149) = 13.95$, $p < .001$, and the pretest and the delayed posttest, $F(1,149) = 11.37$, $p = .001$. Bonferroni correct pairwise comparisons revealed that high prior knowledge students had non-significantly different gains scores between the two conditions (pretest-immediate $p = .96$; pretest-delayed $p = .94$). In contrast, low prior knowledge students had significantly higher gain scores in the Game condition compared to the Non-Game condition (pretest-immediate $p < .001$; pretest-delayed $p < .001$). These results suggest that while high prior knowledge students learned about the same amount from both conditions, students who were categorized as being low prior knowledge students benefitted more from the Decimal Point game than the traditional problem-solving environment. It is of interest to note that this is in line with the "expertise reversal effect" (Kalyuga, 2007), which has been demonstrated in a variety of learning science studies of worked examples, split attention, and a variety of multimedia learning materials. In these studies, an instructional technique that benefits low prior knowledge learners loses its benefit for, and in some cases harms, high prior knowledge learners.

DISCUSSION

In contrast to relatively scant prior evidence that learning mathematics with an educational game is better than a conventional approach (Mayer, 2014), the results of this study show a clear benefit to learning mathematics with an educational game. Students in the Game group learned significantly more, enjoyed their experience more, and made fewer errors during the intervention than the students in the Non-Game group.

We were expecting that one cost of the increased enjoyment for the Game participants would be that the students would take longer to complete the intervention than the Non-Game students, given the additional game steps and visual details the students were presented with in the Game group. However, this did not happen; the Game participants took longer than the Non-Game participants, but the difference was not significant. The significantly higher number of errors made by the Non-Game participants likely explains why it took them longer than expected to complete the materials. If the Non-Game participants were bored or unmotivated, this may have led them to make more errors and, in turn, enjoy their experience less.

Another important finding is that low prior knowledge students learned more about decimals from game playing than from using more conventional learning materials. In fact, the low prior knowledge students may be precisely the best targets for mathematics educational games. These students perform less well and seem to struggle more with mathematics, perhaps because they lack self-motivation or interest. Games might get such students more excited and engaged in mathematics learning. This study provides some real evidence for this often-cited (but not often substantiated) claim.

There are several potential issues with the results that we have carefully addressed and discuss here. First, the students were not randomly assigned to condition, due to the practical constraints of conducting the study in a school, as opposed to a laboratory. Because students might have been distracted by seeing fellow students working with potentially more (or less) interesting materials, we decided to assign entire classes, rather than individual students, to condition. However, we addressed this issue by (a) balancing the classes between conditions according to teacher-reported performance level and (b) confirming that the students in the two conditions were relatively balanced in terms of gender and age. Second, we found that the conditions were significantly different on the pretest, with the Game group having a significantly higher mean score. To address this issue we used appropriate statistical techniques (MANCOVA, ANCOVA, using pretest performance as a covariate) to evaluate the learning benefits and account for the varying pretest results. It should also be noted that higher pretest scores for the Game group could be seen as more of a disadvantage than an advantage for that group, since they had less room to improve. Nevertheless, the Game group gained significantly more than the Non-Game group from pretest to posttest and from pretest to delayed posttest. Finally, we had a relatively high dropout rate (52 out of 213 or ~24%) due to students not completing one or more of the sections of the study materials (i.e., absences, students working exceptionally slow and not completing all materials, and other factors beyond our control). This is addressed by the fact that an equal number of students were eliminated from each group (26 from the Game group, 26 from the Non-Game group) and the finding that the pretest scores of the eliminated students were not significantly different across the groups (Game: 18.0; Non-Game: 18.5, $p = .81$).

Of course, important questions still to ask are why did the Decimal Point game lead to better learning and why was it more enjoyable than the non-game environment? To answer these questions, it's important first to identify the specific design characteristics of the game. Table 4 provides a summary of the game design of Decimal Point.

As shown in Table 4, the Decimal Point game is a relatively simple game. Students do not receive badges, they don't compete against their classmates, the games are deterministic, not random, the game mechanics are uncomplicated, the game is system controlled (students are coerced, step-by-step, through the amusement park), and the game has a very straightforward narrative. Yet, simple gamification may be just what is needed for learning, as suggested by others (cf. Lee & Hammer, 2011; Sheldon, 2011) and as indicated in the Clark et al. (2015) meta-analysis. The Clark et al. (2015) findings suggest that more complex game designs do not equate to better learning outcomes, as both simple and complex game mechanics have led to positive learning effects. There may actually be an important instructional design advantage to simplicity, by not distracting students with unnecessary decision-making, competition, and seductive details (Harp & Mayer, 1998). Furthermore, Sitzmann (2011) showed that highly entertaining games do not lead to better learning outcomes. Thus, the relatively simple approach of Decimal Point may have been a key reason for its success.

Also related to simplicity, the non-competitive, single-player mini-games of Decimal Point encourage an atmosphere where it is OK to fail which may, in turn, contribute to enjoyment of the game. Students know they are supposed to improve over time as they learn to play the game and are not competing against their fellow students as they try to improve. These characteristics may support self-efficacy, and may be a key reason why the lower prior knowledge students benefitted more. They are more likely to be unsuccessful at the beginning of the game, but learn to play and get a better grasp on decimals as they progress. The meta-analysis of Clark et al., (2015) supports this supposition: They found that single-player games without competition outperformed single-player games with

Table 4. Game design characteristics and Decimal Point's design

Game Design Characteristic[3]	Decimal Point Game Design
Is the game single player or collaborative?	Single player
Does the game have a reward structure? (e.g., badges)	No
Is the game competitive? (i.e., Does it have a leaderboard showing other students' scores?)	No
Does the game have a goal structure?	Yes, but a very simple one. The student must make his/her way all the way through the amusement park, playing each mini-game in turn.
Does the game have random behavior in it?	No. The games are presented to the student in sequence; the decimal problems are fixed.
Does the game have simple or elaborate game mechanics?	Simple game mechanics. Generally, in each mini-game the student can only do a single action, such as shoot a dart, type in a number, or select greater than / less than.
Does the student control the game or does the system dictate what the student will do?	System controlled. Student must play games in a prescribed order.
Does the game have significant feedback/scaffolding to guide the student?	Only limited and basic actions are allowed. Correctness feedback is provided.
Does the game have a narrative?	Yes, the game has a very basic narrative. The student makes his/her way through the amusement park, encouraged by the fantasy characters.
Does the game have spaced or massed learning?	Spaced learning. The student can stop at any time and restart the game. The game generally takes several class periods to complete.

competition and suggest that this is inline with research on motivation and learning that shows that competition can undermine self-efficacy (e.g., Bandura, 1997; Pintrich, 2003).

The students played the Decimal Point game over multiple class periods, rather than all at once, and this may have also contributed to the game's learning benefits relative to the more conventional learning environment. Researchers have found that spaced versus blocked learning can result in more positive learning outcomes (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; McDaniel, Fadler, & Pashler, 2013), and the meta-analysis of Wouters et al. (2013) found that media comparison studies in which students played a game for more than a single session led to more beneficial learning outcomes than non-game control conditions. Wouters et al. (2013) conjectured that, in comparison to a more conventional learning environment, the benefits of a game "may pay off only after multiple training sessions in which the players get used to the game" (p. 251). Given the Wouters et al. finding, Clark et al. (2015) hypothesized that games involving multiple sessions would lead to significantly better learning than those involving single sessions. Their meta-analysis confirmed this hypothesis.

CONCLUSION

The study and results reported in this paper represent one of the few rigorous studies in the domain of mathematics in which an educational game has been shown to foster superior performance on learning outcomes as compared to a more conventional instructional approach that presents precisely the same mathematics content. The study in particular shows that students who learned with Decimal Point enjoyed their experience more than learning with a more conventional approach and that low prior knowledge students benefited more from the Decimal Point game than high prior knowledge students. Thus, this study represents a key step in moving from the hope of educational games to an evidence-based reality of educational games. In addition, the relative simplicity of the game's design,

i.e., the single-player format, the uncomplicated game mechanics, and the straightforward narrative, as well as the lack of competition and spaced game play may lead to a better understanding of the design considerations integral to the success of computer-based educational games.

Yet, there are also some limitations to bear in mind. First, caution is warranted regarding the result that the Game group enjoyed their experience more than the Non-Game group. This could be a novelty effect, since the more traditional approach to learning mathematics experienced by most of these students – class lectures, followed by in-class and homework problems to solve – is likely not as novel or engaging as learning from a game. On the other hand, most of the students in this study had likely used computer-based games for learning mathematics before this study (as reported by the teachers). To tease out the novelty effect we would need to conduct a study over a longer period of time, testing whether the enjoyment found in this relatively short-term study persists over time. Second, it should be noted that Decimal Point is somewhat narrowly focused on the single domain of decimals. Finally, as with the vast majority of educational technology research studies, that the study was conducted with a relatively small, specific population of students in a single city of the U.S. Thus, the results may not generalize to other domains or to other populations of students.

To better understand the effects of our game at a fine-grain level we are first planning to conduct a replication study at a new middle school. We hope to run the study with a sample of students comparable to the current study and, to address the first limitation cited above, to run the study for a longer period of time. We will also analyze student log data from the new study, as well as from the study presented in this paper, by applying educational data mining (EDM) techniques to explore the differences between the two media comparison conditions. For instance, we will use moment-by-moment learning models (MBML – Baker, Goldstein, & Heffernan, 2011) to understand whether students learn immediately or in a delayed way from the games (and the conventional materials). In addition, we will analyze learner affect using previously validated automated detectors of confusion and other affective states (Baker et al., 2012). We will use the log analyses to investigate how each of the steps of the student interaction with the games – e.g., game steps and explaining answers – promotes learning and positive affect.

ACKNOWLEDGMENT

REFERENCES

Adams, D., McLaren, B. M., Durkin, K., Mayer, R. E., Rittle-Johnson, B., Isotani, S., & Van Velsen, M. (2014). Using erroneous examples to improve mathematics learning with a web-based tutoring system. *Computers in Human Behavior, 36C (2014),* 401-411. doi:10.1016/j.chb.2014.03.053

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education*, *19*(2), 105–154. doi:10.1109/DIGITEL.2010.55

Aleven, V., Myers, E., Easterday, M., & Ogan, A. (2010). Toward a framework for the analysis and design of educational games.*Proceedings of the 2010 IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning* (pp. 69-76). doi:10.1109/DIGITEL.2010.55

Baker, R. S. J. d., Goldstein, A. B., & Heffernan, N. T. (2011). Detecting learning moment-by-moment. *International Journal of Artificial Intelligence in Education*, *21*(1-2), 5–25.

Baker, R. S. J. d., Gowda, S. M., Wixon, M., Kalka, J., Wagner, A. Z., Salvi, A., & Rossi, L. et al. (2012). Towards Sensor-free Affect Detection in Cognitive Tutor Algebra.*Proceedings of the 5th International Conference on Educational Data Mining* (pp. 126-133).

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Brueckner, L. J. (1928). Analysis of difficulties in decimals. *The Elementary School Journal*, *29*(1), 32–41. doi:10.1086/456173

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380. doi:10.1037/0033-2909.132.3.354 PMID:16719566

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, R., & Glaser, R. (1989). Self explanations: How students study and used examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. doi:10.1207/s15516709cog1302_1

Chi, M. T. H., DeLeeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *25*(4), 471–533. doi:10.1207/s15516709cog2504_1

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2015). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*. doi:10.3102/0034654315582065 PMID:26937054

Clark, R. E. (2001). *Learning from media*. Greenwich, CT: Information Age Publishers.

Conati, C., & Manske, M. (2009). Evaluating adaptive feedback in an educational computer game. *Proceedings of IVA 2009, 9th International Conference on Intelligent Virtual Agents, LNAI* (Vol. 5773, pp. 146-158). Springer Verlag. doi:10.1007/978-3-642-04380-2_18

Forlizzi, J., McLaren, B. M., Ganoe, C. H., McLaren, P. B., Kihumba, G., & Lister, K. (2014). Decimal Point: Designing and developing an educational game to teach decimals to middle school students. In C. Busch (Ed.) *Proceedings of the 8th European Conference on Games Based Learning* (*ECGBL '14*). (pp. 128-135) Academic Conferences and Publishing International Limited, Reading, U.K.

Gamesandlearning.org. (2015). Teachers surveyed on using digital games in class: A games and learning research report. Retrieved from http://www.gamesandlearning.org/2014/06/09/teachers-on-using-games-in-class/#

Gee, J. P. (2003). *What video games have to teach us about learning and literacy* (1st ed.). New York: Palgrave Macmillan.

Glasgow, R., Ragan, G., Fields, W. M., Reys, R., & Wasman, D. (2000). The decimal dilemma. *Teaching Children Mathematics*, *7*(2), 89–93.

Graeber, A., & Tirosh, D. (1988). Multiplication and division involving decimals: Preservice elementary teachers' performance and beliefs. *The Journal of Mathematical Behavior*, *7*, 263–280.

Habgood, M. P. J., & Ainsworth, S. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences*, *20*(2), 169–206. doi:10.1080/10 508406.2010.508029

Hannifin, R. D., & Vermillion, J. R. (2008). Technology in the classroom. In T. L. Good (Ed.), 21st Century education: A reference handbook (Vol. 2, pp. 209–218). Thousand Oaks, CA: SAGE. doi:10.4135/9781412964012. n73

Harp, S., & Mayer, R. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414–434. doi:10.1037/0022-0663.90.3.414

Hiebert, J., & Wearne, D. (1985). A model of students decimal computation procedures. *Cognition and Instruction*, *2*(3-4), 175–205. doi:10.1080/07370008.1985.9648916

Hiebert, J., Wearne, D., & Taber, S. (1991). Fourth graders gradual construction of decimal fractions during instruction using different physical representations. *The Elementary School Journal*, *91*(4), 321–341. doi:10.1086/461658

Honey, M. A., & Hilton, M. L. (2011). Learning science through computer games and simulations. The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=13078&page=R1

Irwin, K. C. (2001). Using everyday knowledge of decimals to enhance understanding. *Journal for Research in Mathematics Education*, *32*(4), 399–420. doi:10.2307/749701

Isotani, S., McLaren, B. M., & Altman, M. (2010). Towards intelligent tutoring with erroneous examples: A taxonomy of decimal misconceptions. In V. Aleven, J. Kay, J. Mostow (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (ITS '10), LNCS (Vol. 6094, pp. 346-348). Berlin: Springer. doi:10.1007/978-3-642-13437-1_66

Johnson, C. I., & Mayer, R. E. (2010). Adding the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior*, *26*(6), 1246–1252. doi:10.1016/j.chb.2010.03.025

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*(4), 509–539. doi:10.1007/s10648-007-9054-3

Ke, F. (2008). A case study of computer gaming for math: Engaged learning from gameplay? *Computers & Education*, *51*(4), 1609–1620. doi:10.1016/j.compedu.2008.03.003

Kouba, V. L., Carpenter, T. P., & Swafford, J. O. (1989). Number and operations. In M. M. Lindquist (Ed.), *Results from the Fourth Mathematics Assessment of the National Assessment of Educational Progress* (pp. 64–93). Reston, VA: National Council of Teachers of Mathematics, Inc.

Lee, J. J., & Hammer, J. (2011). Gamification in education: What, how, why bother? *Academic Exchange Quarterly*, *15*(2), 1–5.

Lenhart, A., Kahne, J., Middaugh, E., Macgill, A., Evans, C., & Vitak, J. (2008). Teens, video games, and civics. Pew Research Center. Retrieved from http://www.pewinternet.org/2008/09/16/teens-video-games-and-civics/

Lepper, M. R., & Malone, T. W. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. E. Snow & M. J. Farr (Eds.), Aptitude, learning, and instruction: Vol. 3. Cognitive and affective process analysis (pp. 255-286). Hillsdale, NJ: Erlbaum.

Lomas, J. D., Patel, K., Forlizzi, J., & Koedinger, K. (2013). Optimizing Challenge in an Educational Game Using Large-Scale Design Experiments.*Proceedings of CHI 2013*. New York, NY: ACM Press. doi:10.1145/2470654.2470668

Mayer, R. E. (2014). *Computer games for learning: An evidence-based approach*. Cambridge, MA: MIT Press.

Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research*, *42*(3), 241–265. doi:10.2190/EC.42.3.a

McDaniel, M. A., Fadler, C. L., & Pashler, H. (2013). Effects of spaced versus massed training in function learning. *Journal of Experimental Psychology*, *39*, 1417–1432. doi:10.1037/a0032184 PMID:23565787

McLaren, B. M., Adams, D. M., & Mayer, R. E. (2015). Delayed learning effects with erroneous examples: A study of learning decimals with a web-based tutor. *International Journal of Artificial Intelligence in Education*, *25*(4), 520–542. doi:10.1007/s40593-015-0064-x

McNamara, D. S., Jackson, G. T., & Graesser, A. C. (2010). Intelligent tutoring and games (ITaG). In Y. K. Baek (Ed.), *Gaming for classroom-based learning: Digital role-playing as a motivator of study* (pp. 44–65). Hershey, PA: IGI Global. doi:10.4018/978-1-61520-713-8.ch003

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

O'Neil, H. F., & Perez, R. S. (2008). *Computer games and team and individual learning*. Boston: Elsevier.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, *95*(4), 667–686. doi:10.1037/0022-0663.95.4.667

Prensky, M. (2006). *Don't bother me, Mom, I'm learning! How computer and video games are preparing your kids for 21st century success and how you can help*. St. Paul: Paragon House.

Putt, I. J. (1995). Preservice teachers ordering of decimal numbers: When more is smaller and less is larger! *Focus on Learning Problems in Mathematics*, *17*(3), 1–15.

Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S., & Peled, I. (1989). Conceptual bases of arithmetic errors: The case of decimal fractions. *Journal for Research in Mathematics Education*, *20*(1), 8–27. doi:10.2307/749095

Riconscente, M. M. (2013). Results from a controlled study of the iPad fractions game Motion Math. *Games and Culture*, *8*(4), 186–214. doi:10.1177/1555412013496894

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346–362. doi:10.1037/0022-0663.93.2.346

Sackur-Grisvard, C., & Léonard, F. (1985). Intermediate cognitive organizations in the process of learning a mathematical concept: The order of positive decimal numbers. *Cognition and Instruction*, *2*(2), 157–174. doi:10.1207/s1532690xci0202_3

Schaaf, R. (2012). Does digital game-based learning improve student time-on-task behavior and engagement in comparison to alternative instructional strategies? *Canadian Journal of Action Research*, *13*(1), 50–64.

Sheldon, L. (2011). *The multiplayer classroom: Designing coursework as a game*. Boston, MA: Cengage.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology*, *64*(2), 489–528. doi:10.1111/j.1744-6570.2011.01190.x

Siwek, S. E. (2010). Video Games in the 21st Century. On the *Entertainment Software Association* website. http://www.theesa.com/facts/pdfs/VideoGames21stCentury_2010.pdf

Squire, K., & Jenkins, H. (2003). Harnessing the power of games in education. *Insight (American Society of Ophthalmic Registered Nurses)*, *3*(7), 5–33.

Stacey, K., Helme, S., & Steinle, V. (2001). Confusions between decimals, fractions and negative numbers: A consequence of the mirror as a conceptual metaphor in three different ways. In M. *v. d.* Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* (*Vol. 4*, pp. 217-224). Utrecht: PME.

Tarng, W., & Tsai, W. (2010). The design and analysis of learning effects for a game-based learning system. *World Academy of Science, Engineering and Technology, 61,* 336-345. doi:10.1.1.192.8685

Tobias, S., & Fletcher, J. D. (2011). *Computer games and instruction*. Charlotte, NC: Information Age.

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. doi:10.2190/FLHV-K4WA-WPVQ-H0YM

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. doi:10.1037/a0031311

ENDNOTES

[1]   Five studies in second-language learning, three in language arts, and three in social studies also met the criteria. Of these domains, only second-language learning, with 4 out of 5 studies showing advantages for games over traditional approaches, has thus far shown the promise of educational games.

[2]   While we include the "Perceived Material Difficulty" category here, we did not analyze and do not report this category in our "Results" section because it is not a focus of any of our research questions and because of significant differences between, for instance, the "material difficulty" and "easily understood assignment" questions.

[3]   There is, of course, no definitive list of "game design characteristics". However, the design characteristics listed in this table are compiled from those suggested by Clark et al. (2015) and Lepper and Malone (1987) and are, arguably, a relatively comprehensive list.