CRESST REPORT 797

Rebecca E. Buschang

Gregory K. W. K Chung

Jinok Kim

AN EXPLORATORY STUDY OF THE RELATIONSHIP BETWEEN COLLABORATION AND MATHEMATICS AND GAME OUTCOMES

JUNE, 2011

The National Center for Research on Evaluation, Standards, and Student Testing

Graduate School of Education & Information Sciences
UCLA | University of California, Los Angeles

An Exploratory Study of the Relationship Between Collaboration and Mathematics and Game Outcomes

CRESST Report 797

Rebecca E. Buschang, Gregory K. W. K Chung, and Jinok Kim
CRESST/University of California, Los Angeles

June, 2011

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
300 Charles E. Young Drive North
GSE&IS Bldg., Box 951522
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305C080015.

The findings and opinions expressed here do not necessarily reflect the positions or policies of the National Center for Education Research, the Institute of Education Sciences, or the U.S. Department of Education.

To cite from this report, please use the following as your APA reference:
Buschang, R. E., Chung, G. K. W. K., & Kim, J. (2011). *An exploratory study of the relationship between collaboration and mathematics and game outcomes.* (CRESST Report 797). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
.

TABLE OF CONTENTS

AN EXPLORATORY STUDY OF THE RELATIONSHIP BETWEEN COLLABORATION AND MATHEMATICS AND GAME OUTCOMES

Rebecca E. Buschang, Gregory K. W. K. Chung, and Jinok Kim
CRESST/University of California, Los Angeles

Abstract

This study is an exploratory study of the relationship between collaboration and mathematics and game outcomes in a video game aimed at teaching concepts related to rational numbers. The sample included 243 middle school students who played the video game either with one partner or individually for 40 minutes. Results suggest that participants with high and low prior knowledge benefit from different conditions during gameplay. Participants with low prior knowledge tend to perform better on math outcomes by working collaboratively and participants with high prior knowledge tend to perform better on math outcomes by working individually. These results are similar to prior findings from classroom research which indicates that collaboration is more effective for low-performing students. These results have implications for designing game environments for low-performing students.

Introduction

The use of video games in classrooms has gained support in recent years. Proponents of the use of video games in classrooms assert that video games provide students with complex, challenging, and situated learning environments that build skills, knowledge, and habits that are difficult to provide in traditional classrooms (Gee, 2003, 2004; Prensky, 2003; Shaffer, Squire, Halverson, & Gee, 2005).

Video game research in the classroom has spanned a wide range of topics, including social studies, with the use of commercial games such as *Civilization* to teach secondary school students (Squire, 2005), and algebra, with games such as *Lure of the Labyrinth* which aimed to teach specific math concepts to middle school students as part of a curricular unit (Lure of the Labyrinth, 2007). Despite the wide range of topics studied, research on K-12 instructional video games has most often been conducted with individually played games. It was not until recently that multiple player K-12 instructional video games have been researched. However, these studies typically investigate content learning or aspects of motivation in games where all students work collaboratively to solve problems or interact with others during simulations (Dieterle, 2009; Ketelhut, Dede, Clarke, & Nelson, 2006; Neulight, Kafai, Kao, Foley, & Galas, 2007; Rosenbaum, Klopfer, & Perry, 2007). One of the few studies that examined collaboration as a variable investigated competitive gameplay, collaborative gameplay, and students working on worksheets individually instead of playing the video game to investigate attitudes towards math and math learning. By comparing the different treatments, researchers found a significant main effect for gameplay on posttest scores and that collaborative gameplay was more effective than the other conditions at promoting positive attitudes towards math (Ke & Grabowski, 2007).

In the classroom, collaborative learning is a widely used strategy and an extensively studied topic. Overall, collaborative grouping has been found to have a positive effect on learning compared to whole group instruction (Abrami, Lou, Chambers, Poulsen, & Spence, 2000; Slavin, 1990, 1991). Therefore, the blend of collaborative gameplay with educational video games warrants investigation.

The purpose of this study was to explore collaborative gameplay with a K-12 instructional video game. To explore this topic, a prototype of a video game aimed at teaching participants aspects of fractions was developed. This game was used with groups of two students paired homogeneously on math achievement.

Literature Review

Collaboration in Classrooms

The research on collaborative learning environments (also called cooperative learning in classroom research) is extensive and includes studies from all major academic subjects and grade levels. In general, reviews of the literature report positive math outcomes for collaborative learning over whole class instruction (Abrami et al., 2000; Slavin, 1990, 1991). One reason for higher math outcomes is that cooperative learning provides more frequent opportunities for student interaction over whole group instruction (Shachar & Sharan, 1994). These interactions, such as frequent help giving and receiving, allow students to verbalize their learning, resolve cognitive conflicts, and learn from peers, and such processes have been reported more frequently with high-performing students than low-performing students (Webb, 1982).

Factors Influencing Collaborative Groupings

Higher posttest scores have been linked to specific group compositions (Lou, Abrami, & d'Apollonia, 2001; Lou, Abrami, & Spence, 2000; Lou, Abrami, Spence, Poulsen, Chambers, & d'Apollonia, 1996; Webb, 1982). One grouping factor that has been explored extensively is ability groupings. Results of a meta-analysis have found that low-performing students gain more from heterogeneous groupings, that medium-performing students gain more from homogeneous groupings, and that group composition made no difference for high-performing students (Lou et al., 1996). In a review of the literature on collaboration with technology (simulations, tutoring programs, Internet use, Logo, and hypertext) that analyzed the impact of group size and type of activity on learning outcomes, results indicated that pairs were more beneficial to students than larger groups and that students benefit from collaborative groups more with programs that focus on tutorials and drill and practice rather than exploratory tasks (Lou et al., 2001). While these results do not exhaust the grouping factors that have been explored, they suggest that ability grouping benefits are dependent upon how pairs are made, that pairs are more beneficial than larger groups, and that guided tasks are more beneficial than exploratory tasks.

Rationale for Cooperative Video Gameplay

In video games, collaboration has not become mainstream practice until recently. With games such as *World of Warcraft* or some Wii games, true cooperation is necessary. However, while there are exceptions, most video games are played individually or competitively against others either virtually with features such as high score boards or in person against others playing in the same room. Cooperative games are different from competitive games because they are those where “all participants work together as a team, sharing the payoffs and outcomes; if the team wins or loses everyone wins or loses” (Zagal, Rick, & His, 2006).

And while the research suggests a general positive effect of collaboration on learning, the use of collaboration in educational games is uncommon with few empirical studies being conducted on the impact of collaboration on math outcomes. One such study investigated how a math game influenced elementary students’ math performance and attitudes toward math. One hundred twenty five students either played a cooperative game, played a competitive game, or were placed in the control condition where students completed drill worksheets over a four-week period (Ke & Grabowski, 2007). Results indicate that gameplay was more effective than drills at increasing math performance, but cooperative and competitive gameplay were equally as effective at promoting math performance. Cooperative gameplay was more effective than competitive gameplay and drills at promoting positive attitudes towards math. Similarly, cooperative gameplay was found to be more effective at promoting higher performance on posttest scores and attitudes towards the technology than competitive gameplay or individual gameplay using a map-reading task (Johnson & Johnson, 1996). These studies suggest that when technology is incorporated in the learning environment, there is a positive impact on learning and attitudes.

Purpose of Research

The purpose of this study was to explore the relationship between homogeneous collaborative grouping and math and game outcomes in a video game setting. Collaboration in this study was defined as two players playing one video game together at one computer. The specific research questions addressed are:

1. Do participants working collaboratively outperform participants working individually on math outcomes?
2. Do participants working collaboratively outperform participants working individually on game outcomes?

Method

Research Design

Teachers were recruited for a larger project through recruitment letters distributed to all mathematics teachers at schools who agreed to participate. Interested teachers contacted the researcher, received more information about the study details, and were placed on a participation list. Two schools were selected for studies testing different aspects of feedback on student math and game performance. Participant data from these studies were combined for the current study for exploratory analyses of the relationship between collaboration and math and game outcomes during video gameplay. Because the study design contained only one school in each condition, classroom characteristics may be confounded. However, this study does not claim to make causal inferences, but instead only uses the data for exploratory purposes.

Students at one school played the video game collaboratively (two students at one computer). Collaborative student pairs were formed by matching two students with similar weekly gameplay experience and pretest scores. Students at the other school played the video game individually (one student at one computer). The content of the video game, gameplay time, and pre- and posttest items were identical between the two conditions. Participants in both conditions also completed survey items individually. Data from a total of 243 middle school students from the two participating schools were analyzed, resulting in 110 participants in the collaborative gameplay condition and 133 participants in the individual gameplay condition.

All students and teachers were informed they could withdraw from the study at any point. No participants withdrew from the study. Teachers were paid $100 per participating class.

Game Description

All participants played a video game, called *Save Patch*, aimed at teaching key topics of rational numbers including: identification of unit size, the numerator, the denominator, and addition of fractions. *Save Patch* is a prototype game that was used as a research test bed (see Figure 1). In the rational number addition video game, participants are presented with the challenge of bouncing a small sack-like doll over various hazards to get it safely to the other side. To do so, participants place small trampolines at various fixed locations along a one- or two-dimensional grid. Each trampoline is made “bouncy” by dragging coils onto the trampoline. The distance each coil will cause Patch to bounce is commensurate with its length. Therefore, if you add a coil of one unit to a trampoline, that trampoline will cause Patch to bounce exactly one unit. In *Save Patch*, one whole unit is always the distance between two lines. It is this unit that becomes the referent for coils of fractional bounce later on. Coils can be added to a trampoline to increase the distance Patch will bounce; however, only identical coils can be added together. While any size coil can be placed on the trampoline initially, subsequent coils can only be added to the trampoline if they are the same size. Initially, whole unit (integer) coils can be added one at a time, reinforcing the meaning of addition with integers.

The game exploits the fact that real numbers can be broken into smaller, identical parts to facilitate addition and that this process is similar in both integer and fractional addition. The intent is to make explicit connections between integer addition (with which many students have confidence) and fractional addition (with which many students struggle). Moreover, the gameplay requires that players (participants) be attentive to the size of the unit they are adding. Unlike many previous games designed to teach mathematics, however, fluency with the basic ideas (the learning goals as specified in the knowledge specifications) is integral, not ancillary, to gameplay.

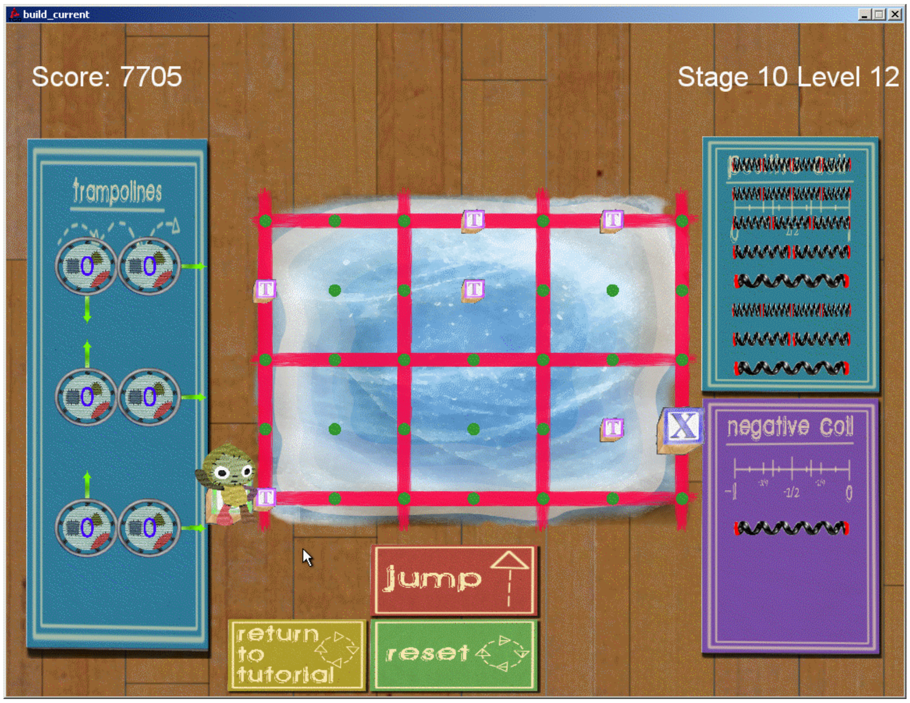


Figure 1. Screenshot of *Save Patch.*

Measures

Math assessments. Pretest and posttest assessments consisted of 18 common math items that addressed the concept of a unit, numerator, denominator, and addition of fractions. One item was adapted from *How Students Learn*: *History, Mathematics, and Science in the Classroom* (National Research Council, 2005) and one from a pre-algebra curriculum called *Powersource* (Phelan, Choi, Vendlinski, Baker, & Herman, 2009; Phelan, Kang, Niemi, Vendlinski, & Choi, 2009). The remaining 16 items were developed specifically for the study. These items were intended to assess a range of knowledge from conceptual understanding to procedural knowledge of the topic. The math assessment items are in Appendix A.

All items had a maximum score of one point. Seventeen items were scored as correct (1 point) or incorrect (0 points), and one item was scored using a rubric of high (1 point), medium (0.66 points), or low (0.33 points). Answers of "don't know" were recoded as incorrect answers. The total score for the pretest and posttest was calculated by summing the points on each item. Reliability using Cronbach's alpha for all scales are reported in Table 1.

Table 1

Reliability of Math Scales (*N* = 243)

	Cronbach's alpha
Pretest	.81
Posttest	.81

Game outcomes. In-game performance was tracked using clickstream data. Clickstream data logged individual actions during gameplay and allowed for tracking of the last level reached, addition errors (e.g., attempting to add unlike fractions), and solution errors (e.g., *Save Patch* deaths).

Weekly gameplay survey. Amount of weekly gameplay was measured via self-report. Participants were asked to respond to the following question, "How many hours a week do you play video games?" using a 5-point scale (1 = *0 hours per week*, 2 = *1-4 hours per week*, 3 = *5-8 hours per week*, 4 = *9-12 hours per week*, 5 = *more than 13 hours per week*).

Participants who worked collaboratively on the game completed this item before playing the game. This information, along with pretest data, was used to form matched pairs among participants who worked collaboratively (see Data Collection Procedure). Participants who played the game individually answered this item after playing the game.

Background survey. Participants' general background information was gathered including age, gender, prior grades in math, and language use at home. The set of background questions is shown in Appendix B. All participants completed these survey items individually after playing the game.

Data Collection Procedure

Due to differences in class period length at the two schools, data collection procedures were slightly different between conditions, but the actual gameplay time remained the same (see Table 2). The main differences were the timing of pretest administration and the number of survey items administered to each condition.

Participants who played the game collaboratively participated for 75 minutes over two class periods on consecutive days. On the first day, a 15-minute pretest and survey were administered individually. This information was used to make homogeneous groups based on both prior game experience and math knowledge. To make pairs, participants were first placed in either the high (5 or more hours per week) or low (less than 5 hours per week) weekly gameplay group based on the game experience question. Next, students' pretest scores were used to classify students as either high math prior knowledge or low math prior knowledge using the mean score as the cutoff point for the groups. Students were then paired with another student with a similar pretest score. The result of this process were pairs of students that fell into one of four categories: (a) high weekly gameplay, high math prior knowledge, (b) high weekly gameplay, low prior math knowledge, (c) low weekly game play, high math prior knowledge, and (d) high weekly gameplay, low prior math knowledge. On the second day, participants played *Save Patch* for 40 minutes with their assigned partner, then completed the posttest and remaining surveys individually.

Participants who played the game individually participated for 75 minutes during one block class period. Participants were first administered the 15-minute pretest individually, then played *Save Patch* for 40 minutes individually, and were then administered the posttest and survey items individually.

Table 2

Administration Procedure

	Collaborative (n = 110)		Individual (n = 133)	
	Timeline	Content	Timeline	Content
Pretest (15 minutes)	Day before gameplay	Math pretest Weekly gameplay survey	Same day as gameplay	Math pretest No survey items
Gameplay (40 minutes)	Day after pretest	*Save Patch*	Same day as pretest	*Save Patch*
Posttest (20 minutes)	Immediately after gameplay	Math posttest Background survey Collaboration survey	Immediately after gameplay	Math posttest Background survey Weekly gameplay survey Game experience survey

Methods

Prior to data analysis, we examined missing data, the shape of the distribution of the various measures, and the equivalence of preexisting characteristics between conditions.

Missing data. Missing data analyses were conducted. Thirteen percent of the sample (28 cases) was dropped due to missing data. Removal of these cases did not affect the mean or variance of the sample. See Appendix C for a full summary of the missing data analysis.

Condition equivalence of preexisting characteristics. Tests for the equivalence of preexisting characteristics between the two conditions were conducted (see Table 3). There were no differences between the conditions on gender, the distribution of self-reported math grades, their self-reported interest in math, or languages spoken at home. However, participants working individually scored significantly higher (M = 12.95, SD = 2.81) on the pretest than participants in the collaboration condition (M = 10.98, SD = 3.51), $t(210) = 4.44$, $p < .001$. There were also slight differences in weekly amounts of gameplay, with participants in the collaborative gameplay condition reporting higher amounts of weekly gameplay. The two conditions were also dissimilar on variations in ethnicity with participants in the collaborative gameplay condition being predominately Asian/Pacific Islander and White/non-Hispanic and participants in the individual gameplay condition being predominately African American and Hispanic/Latina/o.

Analysis of condition equivalence helped us identify two key student characteristics related to the outcomes of the study: weekly gameplay and prior knowledge. We controlled for these important student characteristics in all regression analyses. Thereby, we attempted to account for potential differences between the two non-randomly assigned conditions.

Table 3

Demographic Information

	Collaborative gameplay (n = 110)	Individual gameplay (n = 133)
Grade level		
6th	100%	24%
7th	0%	7%
8th	0%	69%
Gender		
Male	42%	46%
Female	58%	54%
Ethnicity		
Biracial/multiethnic	11%	10%
African American	8%	19%
Asian/Pacific Islander	28%	9%
Hispanic/Latina/o	8%	39%
White, non-Hispanic	39%	11%
Other	7%	11%

Frequency speaking language other than English at home		
Never	38%	27%
Mostly English	37%	31%
Half English, half other	13%	17%
Always other language	13%	25%
Weekly gameplay		
Never	9%	15%
1-4 hours/week	36%	48%
5-8 hours/week	26%	19%
9-12 hours/week	8%	8%
13+ hours/week	20%	9%

Analysis methods. To determine if participants in the collaboration condition outperformed participants working individually on math outcomes, a multiple regression analysis was conducted. The posttest score was modeled as a function of treatment condition controlling for math prior knowledge and weekly amounts of gameplay. The interaction of condition on posttest was also examined using multiple regression analysis. Specifically, the product between the pretest and treatment condition controlling for math prior knowledge and weekly amounts of gameplay was included in the regression equation. In addition, pretest scores were transformed to have a mean value of zero by subtracting the mean value from individual scores in order to estimate the interaction coefficient (Cohen, Cohen, West, & Aiken, 2003).

To determine if participants in the collaboration condition outperformed participants working individually on game outcomes, a multiple regression analysis was conducted. The last level attained was modeled as a function of treatment condition controlling for math prior knowledge and weekly amounts of gameplay. The interaction of condition on last level attained was also examined using multiple regression analysis. Specifically, the product between the last level attained and treatment condition controlling for math prior knowledge and weekly amounts of gameplay was included in the regression equation. Transformed pretests were used as covariates in these regression analyses.

Results

Research question: Do participants in the collaboration condition outperform participants working individually on math outcomes?

There was no main effect for treatment, but the analyses showed a significant interaction effect , $b = -.14$, $t(201) = -2.19$, $p < .05$, indicating that participants with lower math pretest scores tend to benefit more on math outcomes from collaboration than working individually, and that participants with higher math pretest scores tend to benefit more on math outcomes from working individually than working collaboratively (see Table 4 and Figure 2).

Table 4

Results of Regression of Posttest Score

Variable	B	SEB	Beta
Constant	12.01	0.26	
Weekly gameplay	-0.03	0.08	-0.01
Pretest score (centered)	1.00	0.05	1.00***
Condition	0.01	0.21	0.00
Condition × pretest score	-0.14	0.07	-0.11*

Note. $R^2 = .83$ ($n = 201$).
*$p < .05$ (two-tailed). ***$p < .001$ (two-tailed).

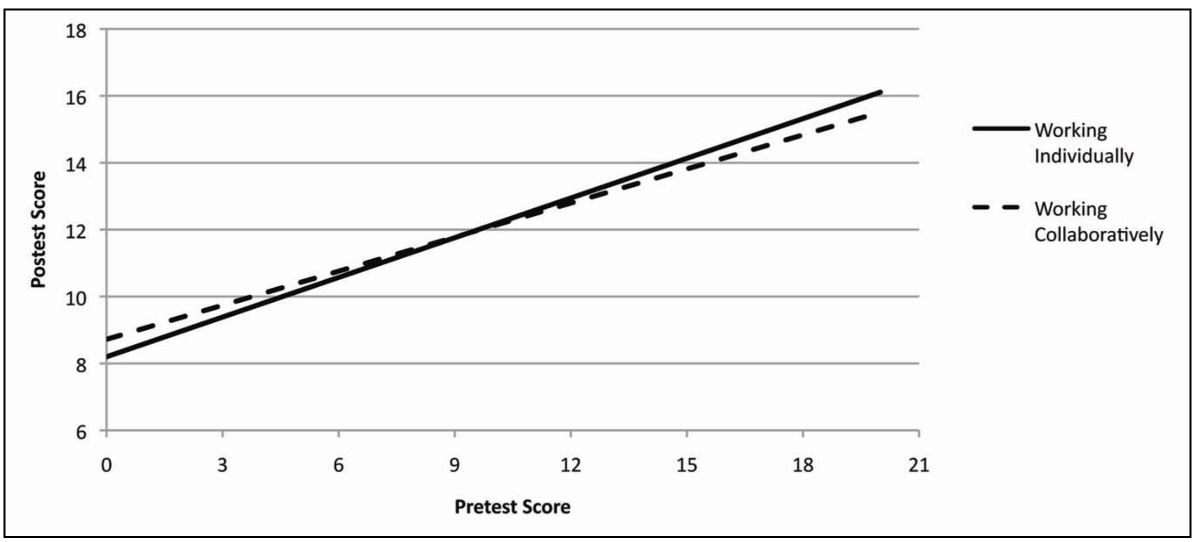


Figure 2. Interaction effect of pretest and condition.

Research question: Do participants in the collaboration condition outperform participants working individually on game outcomes?

There was a main effect for treatment, $p < .001$, indicating that participants who worked individually advanced further in the game than those who worked collaboratively (see Table 5). The analyses also showed a significant interaction effect, $b = -.39$, $t(200) = -2.98$, $p < .01$, indicating that participants with higher math prior knowledge benefited more from working individually and those with lower prior knowledge benefited from working collaboratively (see Figure 3).

Table 5

Results From Regression Analysis of Last Level Attained

Variable	B	SEB	Beta
Constant	18.25	0.75	
Weekly gameplay	.80	.23	.20**
Pretest score (centered)	1.04	0.16	.68***
Condition	-3.10	0.63	-.30***
Condition × pretest score	-.62	0.19	-.32**

Note. $R^2 = .35$ ($n = 211$).
*$p < .05$ (two-tailed). **$p < .01$ (two-tailed). ***$p < .001$ (two-tailed).

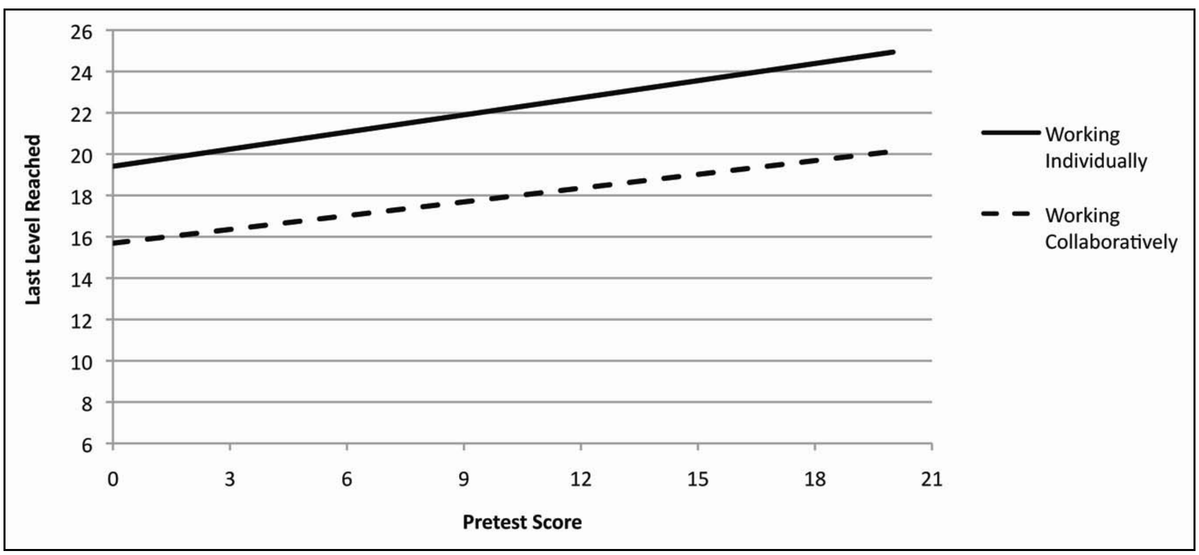


Figure 3. Interaction effect of last level reached by condition and pretest.

Conclusion and Discussion

The results of this exploratory study suggest that treatment condition differentially affected participants with different levels of math prior knowledge. Specifically, participants with low math prior knowledge in the collaboration condition scored higher on posttests, yet did not advance as far in the game as those with low prior knowledge working individually.

Additionally, those with high math prior knowledge in the collaboration condition both scored lower on math outcomes and did not advance as far as those working individually.

During the allotted gameplay time all participants in the collaboration condition spent considerable time explaining how to play the game and the math concepts to each other. The act of explaining a problem to another participant may have helped the explainer's understanding of the problem because explaining a problem often helps an individual internalize problems and generate solutions (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). For the person hearing the explanation, this process may have helped to link prior knowledge to new knowledge about the game and math concepts.

The time to explain the game and math to each other may explain the differential benefit of collaboration for students with high and low math prior knowledge. For students with low math prior knowledge, the opportunity for this explanation was most likely needed due to their low knowledge of the math concepts being presented, and may have helped the pairs with low prior knowledge understand initial game levels and, thus, advance through the game. Students in the individual gameplay condition, however, did not have a partner with whom to clarify the game and generate solutions. This may have led to less understanding of the math concepts being presented and slower advancement through the game. Since more complex math concepts were presented as players advanced through the game, players who did not advance as far were not exposed to the instruction about these math concepts, nor were they able to practice playing levels about these math concepts. Therefore, players who were collaborating during gameplay not only received clarification on the math concepts and how to play the game from their partners, but the explanations may have facilitated understanding of the game and math leading to further advancement in the game. Playing further in the game would, in turn, expose them to more instruction and give them more practice on the math concepts.

For students with high prior knowledge, working individually was more beneficial for students than working collaboratively. One possible reason why working collaboratively was not as beneficial as working individually for students with higher prior knowledge was that students with high prior knowledge did not need the math content explained to them, therefore talking with their partner left less time to play the game. Since the game was set up to become more complex as students advanced through the game, students working collaboratively were not exposed to as much math content as those working individually. Therefore, the instruction and practice in later levels of the game by students working individually may have led to a deeper understanding of the math concepts presented.

The results of this exploratory study suggest that to maximize math outcomes when a limited amount of gameplay is available, participants with low prior knowledge should be grouped with others during gameplay in order to benefit from talking to others about the game, and participants with high prior knowledge should work individually during gameplay to benefit from advancing further in the game.

As noted earlier, this study may be limited by the confounding of school and treatment conditions since we used one intact school in each condition. Key characteristics of students were controlled for in the analysis. Given such limitations and findings, further studies warrant the use of random assignment or clusters (school, classrooms). This will provide more conclusive evidence of the relationship between collaboration and math and game outcomes.

References

Abrami, P. C., Lou, Y., Chambers, B., Poulsen, C., & Spence, J. (2000). Why should we group students within-class for learning? *Educational Research and Evaluation, 6*(2), 158–179. doi:10.1076/1380-3611(200006)6:2;1-E;F158

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182. doi:10.1016/0364-0213(89)90002-5

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.

Dieterle, E. (2009). Neomillennial learning styles and River City. *Children, Youth and Environments, 19*(1), 245–278. Available at http://isites.harvard.edu/fs/docs/icb.topic443490.files/CYE-NLS-and-River-City-Draft2.pdf

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *ACM Computers in Entertainment, 1*(1), 1-4. doi:10.1145/950566.950595

Gee, J. P. (2004). Learning by design: Games as learning machines. *Interactive Educational Multimedia, 8*, 15-23. doi:10.2304/elea.2005.2.1.5

Johnson, D. W., & Johnson, R. T. (1996). Cooperation and the use of technology. In D. H. Jonassen (Ed.), *Handbook of research for educational communications* (pp. 785-811). New York, NY: Macmillan Library Reference.

Ke, F., & Grabowski, B. (2007). Gameplaying for maths learning: Cooperative or not? *British Journal of Education Technology, 38*(2), 249-259. doi:10.1111/j.1467-8535.2006.00593.x

Ketelhut, D. J., Dede, C., Clarke J., & Nelson, B. (2006, April). *A multi-user virtual environment for building higher order inquiry skills in science.* Paper presented at the 2006 AERA Annual Meeting, San Francisco, CA.

Lou, Y., Abrami, P. C., & d’Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*(3), 449–521. doi:10.3102/00346543071003449

Lou, Y., Abrami, P. C., & Spence, J. C. (2000). Effects of within-class grouping on student achievement: An exploratory model. *The Journal of Educational Research, 94*(2), 101–112. doi:10.1080/00220670009598748

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d’Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66*, 423-458. doi:10.3102/00346543066004423

Lure of the Labyrinth. (2007, Version 1). [Online video game]. Learning Games to Go (LG2G).

National Research Council. (2005). *How students learn: History, mathematics, and science in the classroom*. Committee on How People Learn, A Targeted Report for Teachers, M. S. Donovan & J. D. Bransford, Eds. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

Neulight, N., Kafai, Y. B., Kao, L., Foley, B., & Galas, C. (2007). Children's participation in a virtual epidemic in the science classroom: Making connections to natural infectious diseases. *Journal of Science Education and Technology*, *16*(1), 47-58. doi:10.1007/s10956-006-9029-z

Phelan, J., Choi, K., Vendlinski, T. P., Baker, E. L., & Herman, J. L. (2009). *The effects of POWERSOURCE intervention on student understanding of basic mathematical principles* (CRESST Rep. No. 763). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.167.7632&rep=rep1&type=pdf

Phelan, J., Kang, T., Niemi, D. N., Vendlinski, T. P., & Choi, K. (2009). *Some aspects of the technical quality of formative assessments in middle school mathematics* (CRESST Rep. No. 750). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing. Available at http://www.cse.ucla.edu/products/reports/R750.pdf

Prensky, M. (2003). Digital game-based learning. *ACM Computers in Entertainment*, *1*(1), 1-4. doi:10.1145/950566.950596

Rosenbaum, E., Klopfer, E., & Perry, J. (2007). On location learning: Authentic applied science with networked augmented realities. *Journal of Science Education and Technology*, *16*(1), 31-45. doi:10.1007/s10956-006-9036-0

Shachar, H., & Sharan, S. (1994). Talking, relating, and achieving: Effects of cooperative learning and whole-class instruction. *Cognition and Instruction*, *12*, 313-353. doi:10.1207/s1532690xci1204_2

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan, 87*(2), 104-111.

Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research, 60*(3), 471-499. doi:10.3102/00346543060003471

Slavin, R. E. (1991). Synthesis of research on cooperative learning. *Educational Leadership, 48*(5), 71-82.

Squire, K. (2005). *Replaying history: Learning world history through playing Civilization III* (Doctoral Thesis, University of Indiana). Retrieved from website.education.wisc.edu/kdsquire/REPLAYING_HISTORY.doc

Webb, N. M. (1982). Student interaction and learning in small groups. *Review of Educational Research*, *52*(3), 421-445. doi:10.3102/00346543052003421

Zagal, J. P., Rick, J., & His, I. (2006). Collaborative games: Lessons learned from board games. *Simulations and Gaming, 37,* 24-40. doi:10.1177/1046878105282279

Appendix A:

Math Assessment Items

QUESTIONS

Fill in the box with a number that will make the statement true.

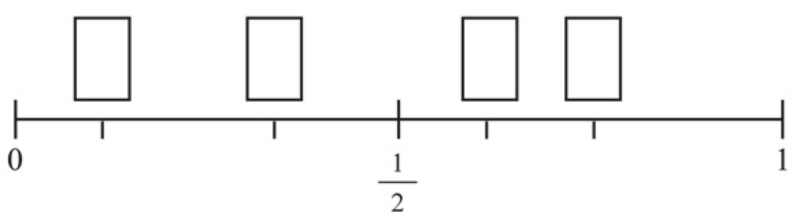
1. $$\frac{2}{10} + \frac{4}{10} + \frac{1}{10} = \frac{7}{\square}$$

2. A student has finished **four and a half pages** of a six-page test. Write a fraction that shows the part of the test the student has completed. Be sure to show your final answer. **Write this as a proper fraction—that is, just whole numbers—no decimals or fractions within fractions.**

   Answer: ____________________

3. Here are four fractions: $\frac{3}{4}$, $\frac{1}{8}$, $\frac{1}{3}$ and $\frac{3}{5}$.

   Look at the number line below. Write each fraction in the correct box.



For the questions below, fill in each box with a number that will make the statement true.
The fractions DO NOT need to be simplified!
Use the chart below to help you add fractions. The chart shows equivalent fractions for $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}$ and $\frac{1}{6}$.

Equivalent Fractions

$\frac{1}{2}$	$=\frac{2}{4}$	$=\frac{3}{6}$	$=\frac{4}{8}$	$=\frac{5}{10}$	$=\frac{6}{12}$	$=\frac{7}{14}$	$=\frac{8}{16}$	$=\frac{9}{18}$	$=\frac{10}{20}$
$\frac{1}{3}$	$=\frac{2}{6}$	$=\frac{3}{9}$	$=\frac{4}{12}$	$=\frac{5}{15}$	$=\frac{6}{18}$	$=\frac{7}{21}$	$=\frac{8}{24}$	$=\frac{9}{27}$	$=\frac{10}{30}$
$\frac{1}{4}$	$=\frac{2}{8}$	$=\frac{3}{12}$	$=\frac{4}{16}$	$=\frac{5}{20}$	$=\frac{6}{24}$	$=\frac{7}{28}$	$=\frac{8}{32}$	$=\frac{9}{36}$	$=\frac{10}{40}$
$\frac{1}{5}$	$=\frac{2}{10}$	$=\frac{3}{15}$	$=\frac{4}{20}$	$=\frac{5}{25}$	$=\frac{6}{30}$	$=\frac{7}{35}$	$=\frac{8}{40}$	$=\frac{9}{45}$	$=\frac{10}{50}$
$\frac{1}{6}$	$=\frac{2}{12}$	$=\frac{3}{18}$	$=\frac{4}{24}$	$=\frac{5}{30}$	$=\frac{6}{36}$	$=\frac{7}{42}$	$=\frac{8}{48}$	$=\frac{9}{54}$	$=\frac{10}{60}$

4. $\frac{1}{5} + \frac{2}{5} = \frac{\square}{\square}$

The fraction does not need to be simplified.

5. $\frac{1}{6} + \frac{1}{4} = \frac{\square}{\square}$

The fraction does not need to be simplified.

6. $\frac{1}{6} + \frac{\square}{\square} = \frac{5}{6}$

The fraction does not need to be simplified.

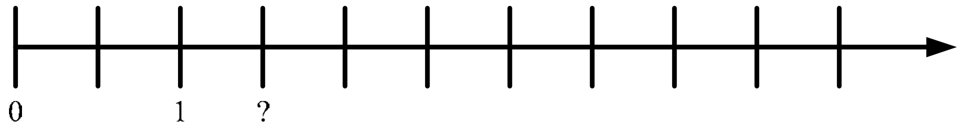
7. $\frac{2}{5} + \frac{3}{10} = \frac{\square}{\square}$

The fraction does not need to be simplified.

8. $\frac{3}{11} + \frac{\square}{\square} = \frac{7}{22}$

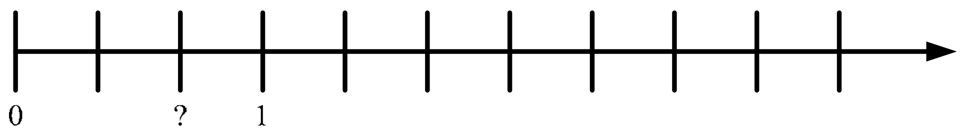
The fraction does not need to be simplified.

9. At what number is the “?” located?



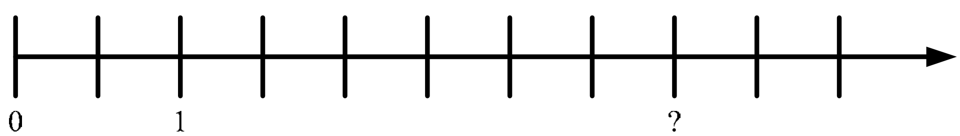
Answer: __________

10. At what number is the “?” located?



Answer: __________

11. At what number is the “?” located?



Answer: __________

Fill in the box with a number that will make the statement true.

12. $\frac{1}{5} + \frac{3}{5} + \frac{\square}{5} = \frac{5}{5}$

13. How many $\frac{1}{4}$'s are in $\frac{3}{4}$?

a. Answer: _________

b. What does the numerator of 3 tell you in $\frac{3}{4}$? **Check only one box.**

- ☐ It tells you there are three $\frac{1}{4}$ in this fraction
- ☐ It tells you the whole unit is broken into three pieces
- ☐ It tells you there are three whole units in this fraction
- ☐ It tells you to add 4 three times

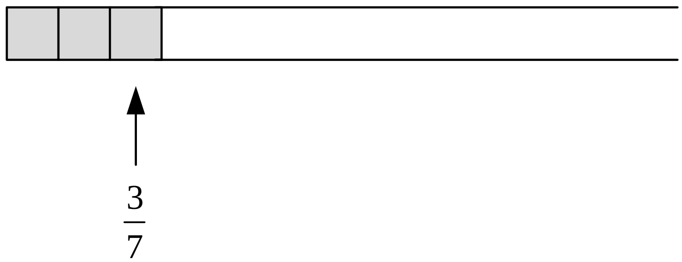
c. What does the denominator of 4 tell you in $\frac{3}{4}$? **Check only one box.**

- ☐ It tells you there are four $\frac{1}{4}$ in this fraction
- ☐ It tells you the whole unit is broken into four pieces
- ☐ It tells you there are four whole units in this fraction
- ☐ It tells you to add 3 four times

Fill in **EACH** box with a number that will make the statement true.

14. $$\frac{1}{3} + \frac{\square}{3} + \frac{\square}{3} + \frac{\square}{3} = \frac{4}{3}$$

15. The figure below shows $\frac{3}{7}$ of a whole unit shaded. Complete the figure to show where the whole unit ends. Be sure to draw lines ("|") to show where each piece is.



Fill in **EACH** box with a number that will make the statement true.

16. $\frac{1}{6} + \frac{1}{6} + \frac{\square}{\square} = \frac{3}{6}$

Appendix B:
Background Questions

BACKGROUND

1. Birth date: ________ / ________
   Month Year

2. Grade: ☐ 4th ☐ 5th ☐ 6th ☐ 7th ☐ 8th ☐ 9th ☐ 10th ☐ 11th ☐ 12th

3. What are you learning in your math class now?
   ____

4. Gender: ☐ Male ☐ Female

5. Ethnicity (choose only one):
   ☐ Biracial / multiethnic
   ☐ African-American
   ☐ Asian or Pacific Islander
   ☐ Hispanic / Latino/a
   ☐ Native-American
   ☐ White, non-Hispanic
   ☐ Other ______________________

6. How often do people in your home talk to each other in a language other than English?
   ☐ Never ☐ Once in a while ☐ About half of the time ☐ All or most of the time

7. What was your math grade on your last report card?
   ☐ A ☐ B ☐ C ☐ D ☐ F ☐ Don’t know

8. What were your math grades last year?
   ☐ A ☐ B ☐ C ☐ D ☐ F ☐ Don’t know

9. Did you play a version of this game before? ☐ Yes ☐ No

Appendix C:
Missing Data Analysis

An examination of participants' missing data responses for the math pretest and posttest items was conducted. Missing was defined as a blank response where there was no indication that the student knew or did not know the answer to the item. Fifteen participants (6.2%) had one or more missing responses on the pretest and 39 participants (16.0% of sample) had one or more missing responses on the posttest (see Table C1). Complete data for the pretest and posttest were available for 212 participants (87% of sample).

Table C1

Number of Missing Responses in Pretest and Posttest Items (*N* = 243)

		No. of missing responses																	
Scale	No. of items	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Pretest	17	204	19	7	7	0	1	0	0	0	0	0	0	0	1	0	0	0	4
Posttest	17	205	20	3	5	1	0	1	0	2	0	0	0	0	0	0	0	0	4

We first examined whether the distribution of the sample with missing pretest and posttest data was different from the original distribution (see Table C2). A chi-square test for the difference in the distribution of participants was not significant ($\chi^2 > 0.00$, $p = .99$), suggesting that the distribution of participants with missing data was not systematically different from the pattern of participants with complete data.

Table C2

Distribution of Missing and Retained Participants by Condition

	Condition		
	Individual (n = 110)	Collaboration (n = 133)	Total
Missing sample	14	17	31
Remaining sample	96	116	212
Total	110	133	243

We chose a maximum missing response rate of 20% for each measure (i.e., allowing at most 3 missing items per math measure), which allowed us to retain 87% of the sample. We then checked whether there were differences in analysis results using the original sample and the reduced sample with respect to distribution of grades, gender, and gameplay experience across conditions. We also checked if the sample characteristics (means and standard deviations) changed significantly before and after dropping cases.

Table C3 shows the means and standard deviations of the full and reduced samples. Independent *t* tests were conducted to check if there were differences between the samples. No significant differences were found on any measure.

Table C3

Comparison of Samples: Mean and Standard Deviations

	Original sample			Sample without missing data		
Measure	*n*	*M*	*SD*	*n*	*M*	*SD*
Pretest	237	11.67	3.34	212	11.87	3.35
Posttest	233	11.99	6.95	212	12.10	3.33

Table C4 shows the skewness and kurtosis of the two samples. Fisher's skewness coefficient was computed for sample for each measure (skewness / standard error of skewness). The sample showed no significant skewness for the original or reduced sample, suggesting the removal of missing cases did not affect the skewness of the distribution. Neither the original nor the reduced sample showed significant kurtosis suggesting neither sample deviated from the normal distribution. Based on these analyses, we concluded that the removal of cases with 3 or more missing values on any measure did not substantially change the shape of the distribution.

Table C4

Comparison of Samples: Skewness and Kurtosis

	Original sample					Sample without missing data				
Measure	*n*	Skewness	*SE*	Kurtosis	*SE*	*n*	Skewness	*SE*	Kurtosis	*SE*
Pretest	237	-0.33	0.16	-0.59	0.31	212	-0.44	0.17	-0.47	0.33
Posttest	233	-0.35	0.16	-0.62	0.32	212	-0.41	0.17	-0.59	0.33

The next set of analyses examined whether dropping the cases with missing data resulted in different distributions across conditions within subgroups with respect to gender, game experience, and self-reported grades in math. As Table C5 shows, the samples were similar with respect to the distribution of participants by gender, weekly gameplay, and self-reported grades.

Table C5

Comparison of Samples: Condition by Gender, Weekly Gameplay, and Self-Reported Grades in Math

	Original sample				Sample without missing data			
	n	Statistic	*df*	*p*	*n*	Statistic	*df*	*p*
Condition × gender	210	0.29	1	0.59	189	0.50	1	0.48
Condition × weekly gameplay	182	1.55	4	0.82	165	0.37	4	0.98
Condition × self-reported math grades	226	9.39	4	0.05	201	7.76	4	0.10

From these analyses we concluded that dropping 31 cases from the original dataset because of missing responses on the math measures did not unduly affect the mean, the variance, the shape of the distribution on the math measures, or the distribution of participants across conditions and various subgroups. Thus, subsequent analyses were based on data from 191 participants.