

# **Effects of Game-Based Learning on Students' Mathematics Achievement: A Meta-Analysis**

## **Abstract**

This report presents findings from a meta-analysis of experimental and quasi-experimental studies investigating effects of instructional games on mathematics achievement of PreK—12<sup>th</sup> grade students compared to traditional classroom methods. School setting (PreK-5<sup>th</sup> vs. 6<sup>th</sup>-12<sup>th</sup>) and type of assessment instrument (research-made vs. standardized) were explored as potential moderators of the relationship between game-based learning and mathematics achievement. Results showed heterogeneity among studies, both in magnitude and direction. Using a random-effects model, a small but marginally significant overall effect ( $d^* = 0.255$ ) suggests that math video games might have contributed to higher learning gains as compared to traditional instructional methods. Furthermore, moderator analyses suggest that this effect does not significantly vary neither between instrument types nor between school settings.

## **Objectives**

The goal of this meta-analysis study was to examine the relationships between instructional game-based interventions for mathematics skills development and student mathematics achievement in PreK-12<sup>th</sup> grades. Specifically, we assessed the relative effectiveness of these interventions as compared to a traditional classroom instruction. Experimental and quasi-experimental studies that compared the effectiveness of video gaming in mathematics with a traditional classroom approach in that content area were collected and analyzed. In addition to overall results that include descriptions and inferences on central tendency, heterogeneity, and related quantities, two moderators (school setting and instrument type) are explored at this time.

## **Perspective**

Learning mathematics presents various challenges for many children. Mathematics is often associated as a difficult and tedious subject to learn (Sedig, 2008). Educational video games have the potential of addressing these challenges. Interactive immersive games can consume children's attention for hours while providing them with effective instruction and an engaging learning experience. Games have been widely used to promote children's mathematics achievement in various domains including problem-solving and algebra skills (Abramovich, 2010), strategic and reasoning abilities (Bottino, Ferlino, Ott & Tavella, 2007), critical geometry skills (Yang & Chen, 2010), and arithmetic procedures (Moreno & Duran, 2004). Nevertheless, National Mathematics Advisory Panel (NMAP, 2008) does not provide a direct recommendation for using games "as a potentially useful tool in introducing and teaching specific subject-matter content to specific populations" (p. 51) due to the limited number of rigorous studies exploring effects of game-based learning on math skills development.

In spite of the high attention that educational games have drawn over the last two decades, empirical findings on the effects of math video gaming on student achievement are inconsistent. For example, Kebritchi (2008) found that high school students that interacted with a math video game *DimensionM* outperformed their non-gaming peers. Okolo (1992) did not find

significant differences between game-based and classroom groups of special needs students; yet, positive results were observed for students with high achievement motivation. A number of studies revealed the opposite effects of math video gaming (e.g., Costabile, De Angeli, Rosclli, Lanzilotti, & Plantamura, 2003). Not only did different research teams that used different games for promoting distinct learning outcomes report mixed results, even findings by the same researcher that used the same math video games were contradictory (e.g., Ke, 2008a, 2008b, 2008c; Ke & Grabowski, 2007).

A meta-analytic review that quantitatively integrates findings of single studies may help gain an understanding of the effectiveness of game-based learning for student math achievement. Several meta-analyses have attempted to synthesize findings of empirical research on game-based learning and academic achievement (e.g., Connolly et al., 2012; Vogel et al., 2006; Young et al., 2012). However, most of these studies were not completed due to methodological challenges associated with the shortage of empirically rigorous research in this area. In addition, these meta-analyses spanned multiple content areas. Because educational video gaming is a developing field that constantly produces new empirical findings and their applications vary across various content areas, we have conducted an extensive search of published and unpublished academic work on video gaming focusing on mathematics content area exclusively two years after the publication of the most recent meta-analytical attempts on instructional gaming (Connolly et al., 2012; Young et al., 2012). Similarly to previous gaming meta-analytical reviews, we used the following definition of a computer game: “A computer game is defined as such by the author, or inferred by the reader because the activity has *goals*, is *interactive*, and is *rewarding* (gives feedback)” (Vogel et al., 2006, p. 231).

## **Inclusion Criteria and Moderator Descriptions**

### **Literature Search**

Online searches of the ERIC, PsycINFO, Wilson, Google Scholar, Jstor, and ISI Web of Science databases were performed to collect studies focusing on the effects of computer games on student mathematics performance published between 2000 through 2014. The following key words were used in order to extract studies for our initial review: *computer games, electronic games, video games, computer software, mathematics achievement, mathematics education, number sense, numerical skills, numbers, experiment, and experimental studies*.

### **Inclusion Criteria**

The initial search located 560 studies. These studies were examined by two reviewers using the following inclusion criteria:

- The study used experimental or quasi-experimental research design
- The study employed game-based and traditional classroom instructional interventions
- Mathematics achievement was used as an outcome
- Study participants were PreK-12<sup>th</sup> grade students
- The study reported sufficient statistical data to calculate effect sizes
- Publication range from 2000-2014

Upon review, 13 studies (providing 17 independent effect sizes) satisfied inclusion criteria to be included in the meta-analysis.

## Selection of Variables

Part of our methods included moderator analyses of select study characteristics. At this point in our research, two study characteristics (school setting and instrument type) were considered. The final version of our paper will include more moderators in effort to explore effect-size heterogeneity as much as possible given our collection of studies.

### *School Settings*

The school settings variable was dichotomous. One level of this factor was named *primary* and consisted of studies (if multiple independent groups were present) with PreK-5<sup>th</sup> grade students (or its equivalent if outside of the United States). The other factor level was named *secondary* and consisted of studies with 6<sup>th</sup>-12<sup>th</sup> grade participants. This variable sought to evaluate whether the effectiveness of mathematics game-based learning varied from one school setting to another to accommodate for a continued increase in difficulty of mathematics skills and decrease in student motivation from preschool-elementary to middle-high school settings (Harter, 1981).

### *Administered instrument type*

The administered instrument type includes also a dichotomous variable. The first factor level, *researcher made*, includes surveys, questionnaires, and tests which were composed by researchers of the study. If researchers used selected questions or portions from a standardized instrument or large-scale assessment, this was then considered to be a researcher-made instrument. The rationale for this is that once altered from the original form, psychometric qualities often begin to fluctuate and the instrument is no longer presented as intended by the original instrument creator(s).

The other factor level, *standardized academic test*, considers utilitarian standardized instrument or large-scale assessments. This type of instruments has long-standing validity, reliability, and psychometric properties which are generally accepted by the research community typically making use of said instruments. The instrument moderator was included to determine whether the type of administered instrument helped explain variability in the effectiveness of game-based learning on student mathematics achievement.

## Quantitative Methods and Results

In order to compare the effectiveness of a game-based learning with a traditional classroom instruction, we calculated Cohen's  $d$  values<sup>1</sup> as the ratio of the mean of a game-based achievement measure and a traditional classroom instruction achievement measure, divided by the pooled standard deviation. Figure 1 shows a forest plot of all 17 effect sizes. The variability of effect-size magnitudes and precision appears somewhat heterogeneous based on interpretations of plotted confidence intervals. Not only do effect-size magnitudes vary in size, their range spans both positive and negative sides of the spectrum; this further suggests a diverse collection of effects. We formally test the suspected effect-size heterogeneity using  $Q$  and  $I^2$  (see Higgins & Thompson, 2002; Higgins et. al. 2003). These statistical assessments,  $Q(16) = 51.88, p < .0001$  and  $I^2 = 69\%$ , confirmed what we supposed from Figure 1 regarding effect-

---

<sup>1</sup> We actually computed a slightly different version of this effect-size metric which is statistically unbiased; see Hedges (1981).

size heterogeneity. Given this result and our generalizability intentions when answering our research questions, we forwent using a restrictive fixed-effect model and made use of random-effects and mixed-effects models to explore effect descriptively and when daring statistical inferences.

We conducted a random-effects model<sup>2</sup> in order to explore the assumed between-studies variability. The overall random-effects<sup>3</sup> weighted effect sizes was  $d^* = 0.255, p = .046$ . The Knapp and Hartung adjusted (Knapp & Hartung, 2003) 95% confidence interval about the mean was [0.005, 0.502]. The overall effect of gaming on mathematical achievement was marginally significant and quite variable, as denoted by the rather wide confidence interval large standard error ( $SE = 0.11$ ). Furthermore, as we noticed in Figure 1, there is reason to believe that collected effects likely vary from study-to-study. With a between-studies variability estimate of  $\tau^2 = 0.14 (SE = .07)$ , we have more reason to believe that the effect of gaming on mathematical achievement is in disagreement among studies. Next we explore possible reasons for this excessive effect variability, beyond sampling error.

At the current point in our research, two study characteristics (i.e., moderators) were explored as possible contributors to effect variability: school level (primary or secondary) and instrument type (researcher made or standardized). Using boxplots, Figure 2 provides distributional characteristics within each group of both moderators. Two mixed-effect models were used, one for each moderator<sup>4</sup>. Both models produced results which described these moderators as not significant factors for explaining between-studies variability. Put another way, the difference in effects between groups in each factor was considered statistically insignificant. That being said, these are quite interesting results. The effect of gaming on mathematical achievement does not seem to vary based on the level of school (primary versus secondary), which is closely related to participants ages. Also, the manner in which achievement is assessed, either using a standardized measure or researcher-made instrument, was statistically irrelevant.

Last, given our limited number of collected studies, we believe it especially important to assess potential publication bias. Figure 3 shows a funnel plot with approximate 90% (white), 95% (light gray), and 99% (dark gray) confidence intervals. The observed effects (filled circles) show a likely asymmetry trend about the point of no effect. This was statistically assessed using the Trim and Fill method (Duval & Tweedie, 2000). Imputed effects are displayed in Figure 3 as unfilled circles. The difference in means (observed effects versus observed and imputed effects combined) was approximately 0.08 standard deviations. Given the visible asymmetry and Trim and Fill result, we believe there is likely some publication bias associated with our collection of studies.

## Discussion and Study Significance

To the best of our knowledge, this report is the first documented effort that synthesizes empirical findings focusing specifically on video gaming in mathematics. The empirical research on games in mathematics is still very limited (Connolly et al., 2012) and our present study has further confirmed the paucity of research in this area. Despite a considerably large number of

---

<sup>2</sup> All analyses and graphics were conducted in R (R Core Team, 2014) using the *metaphor* package (Viechtbauer 2010a; 2010b).

<sup>3</sup> The random-effects model was estimated using restricted maximum likelihood estimation.

<sup>4</sup> We have no suspicion or theoretical justification for an interaction between these two factors.

reviewed studies (above 500), only 13 studies that compared game-based learning with traditional instructional methods were selected (Table 1). In order to make generalizations beyond the collection of studies in this meta-analysis, random-effect and mixed-effect models were utilized. Also, a check for and brief discussion of the presence of publication bias was provided.

Using a random-effects model, a small but marginally significant overall effect ( $d^* = 0.255$ ) suggests that math video games might have contributed to higher learning gains as compared to traditional instructional methods. Furthermore, moderator analyses suggest that this effect does not significantly vary neither between instrument types nor between school settings.

One of the possible explanations of the excessive effect variability revealed in the present meta-analysis could be the game mechanics. However, many researchers did not use a single game as an instructional intervention but a series of games. This limits our ability to determine how a specific game design/genre has influenced the learning process.

Another factor that can possibly explain the relationships between math video gaming and academic achievement is the skills and knowledge promoted in a game. Mathematics includes a wide variety of distinct skills and knowledge that range from basic mathematics skills, to geometry, mathematics word problems, complex computations, and higher level thinking tasks. Examining how games facilitate acquisition of various skills can possibly explain the meta-analysis findings. Therefore, our future steps related to this research include examining whether learning tasks targeted in the select studies and time spent on video gaming can explain the relationships between math video gaming and student achievement.

## References

Studies with an \* were used in the meta-analysis

- Abramovich, S. (2010). *Topics in mathematics for elementary teachers: A technology-enhanced experiential approach*. Charlotte, NC: Information Age Publishing, Inc. (Review by David Fowler.).
- \*Bottino, R.M., Ferlino, L., Ott M. and Tavella, M. (2007). Developing strategic and reasoning abilities with computer games at primary school level, *Computers & Education*, 49 (4) (2007), 1272–1286.
- Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59, 661–686.
- Costabile, M., De Angeli, A., Roselli, T., Lanzilotti, R., & Plantamura, P. (2003). Evaluating the educational impact of a tutoring hypermedia for children. *Information Technology in Childhood Education Annual*. 15(1), 289-308.
- \*Delacruz. (2010). *Games as formative assessment environments: Examining the impact of explanations of scoring and incentives on math learning, game performance, and help seeking*. (PhD Dissertation), University of California.
- \*Din, F. S., & Caleo, J. (2000). *Playing computer games versus better learning*. Paper presented at the Eastern Educational Research Association, Clearwater, FL.
- Duval, S., & Tweedie, R. (2000). A nonparametric ‘Trim and Fill’ method of assessing publication bias in meta-analysis. *Journal of the American Statistical Association*, 95(449), 89-98.
- Fu, Kuang (1999). *My first math adventure – Counting and classification [Computer software]*. Taipei: Kuang Fu Publishing Co.
- \*Gelman, A. (2010). *Mario math with millennials: the impact of playing the Nintendo DS on student achievement*. (PhD Dissertation), University of Denver.
- Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology*, 17, 300-312.
- \*Hawkins, D. (2008). *The application of entertainment video games in elementary mathematics instruction*. (PhD Dissertation), Argosy University.
- Hedges, L. V. (1981). Distribution theory for Glass’ estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107-128.

- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in meta-analysis. *Statistics in Medicine*, *21*, 1539-1558.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analysis. *British Medical Journal*, *327*, 557-560.
- \*Ke, F. (2006). *Computer-Based playing within alternative classroom goal structures on fifth-grades' math learning outcomes: cognitive, metacognitive, and affective evaluation and interpretation*. (Doctoral dissertation) Retrieved from ProQuest Dissertations and Theses. (UMI Number 3229015).
- Ke, F. (2008a). Alternative goal structures for computer game-based learning. *Computer-Supported Collaborative Learning*, *3*, 429–445. doi:10.1007/s11412-008-9048-2
- \*Ke, F. (2008b). A case study of computer gaming for math: Engaged learning from gameplay? *Computer & Education* *51*(2008), 1609-1620
- Ke, F. (2008c). Computer games application within alternative classroom goal structures: Cognitive, metacognitive, and affective evaluation. *Education Tech Research and Development*, *56*, 539–556. doi:10.1007/s11423-008-9086-5
- \*Ke, F., & Grabowski, B. (2007). Gameplaying for maths learning: cooperative or not? *British Journal of Educational Technology*, *38*(2), 249-259. doi: 10.1111/j.1467-8535.2006.00593.x
- Kebritchi, M. (2008). Effects of a computer game on mathematics achievement and class motivation: An experimental study. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *69*(6-A), 2121.
- \*Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education*, *55*(2), 427–443. doi:10.1016/j.compedu.2010.02.007
- \*King, A. (2011). *Using interactive games to improve math achievement among middle school students in need of remediation*. (PhD Dissertation), George Washington University.
- Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, *22*(17), 2693-2710.
- Moreno, R., & Duran, R. (2004). Do multiple representations need explanations? The role of verbal guidance and individual differences in multimedia mathematics learning. *Journal of Educational Psychology*, *96*(3), 492-503.

- National Mathematics Advisory Panel (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*, U.S. Department of Education: Washington, DC.
- Okolo, C. M. (1992). The effect of computer-assisted instruction format and initial attitude on the arithmetic facts proficiency and continuing motivation of students with learning disabilities. *Exceptionality*, 3, 195–211.
- R Core Team. (2014). R: A language and environment for statistical computing (version 3.1.0). Vienna, Austria: R Foundation for Statistical Computing. Available from <http://www.R-project.org>
- \*Sedig, K. (2008). From play to thoughtful learning: A design strategy to engage children with mathematical representations. *Journal of Computers in Mathematics and Science Teaching* (2008)27(1), 65-101.
- SMEC (1999). Toby' IQ training camp – The seed of logic [Computer Software]. Taipei: SMEC Media Entertainment Corp.
- \*Sung, Y.-T., Chang, K.-E., & Lee, M.-D. (2008). Designing multimedia games for young children's taxonomic concept development. *Computers & Education*, 50(3), 1037-1051. doi: 10.1016/j.compedu.2006.07.011
- \*Van Eck, R., & Dempsey, J. (2002). The effect of competition and contextualized advisement on the transfer of mathematics skills a computer-based instructional simulation game. *Educational Technology Research and Development*, 50(3), 23-41. doi: 10.1007/bf02505023
- Viechtbauer, W. (2010a). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48.
- Viechtbauer, W. (2010b). metafor (version 1.6-0) [R]. Available from <http://cran.r-project.org/web/packages/metafor/index.html>
- 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(4), 229-243.
- Yang, J.C. & Chen, S.Y. (2010). Effects of gender differences and spatial abilities within a digital pentominoes game. *Computers & Education*, 55(3), 1220-1233.
- Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., et al. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research*, 82, 61-89.

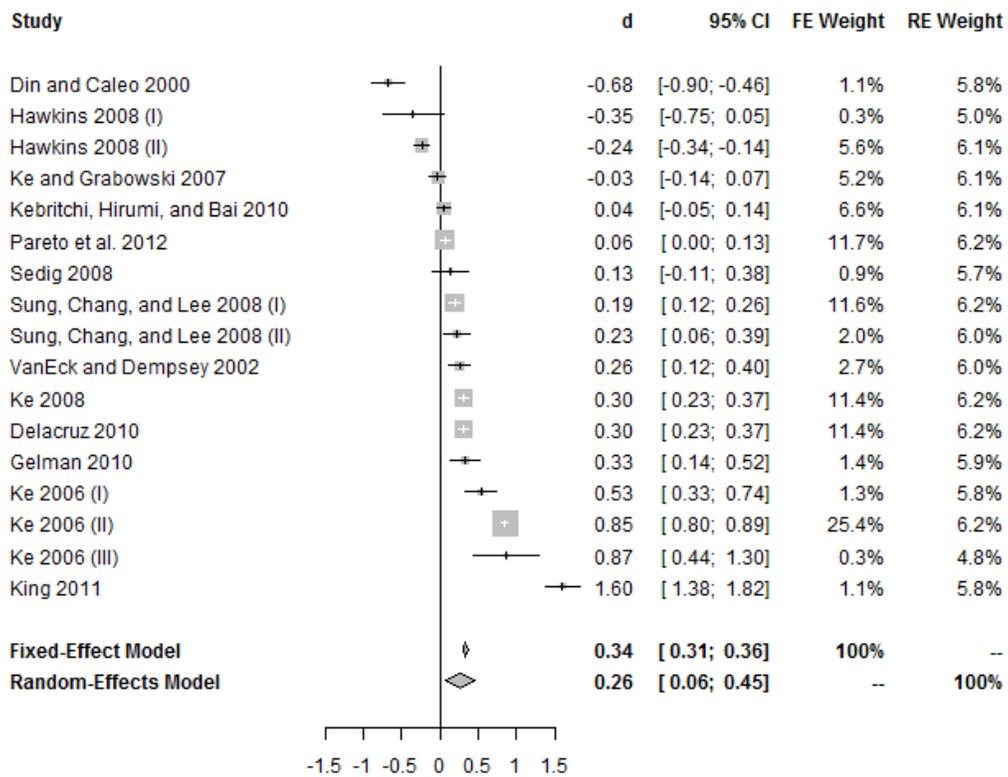


Figure 1. Forest plot for effect sizes. Both fixed-effect and random-effects results are provided. Confidence interval (CI); Fixed Effect (FE); Random Effect (RE).

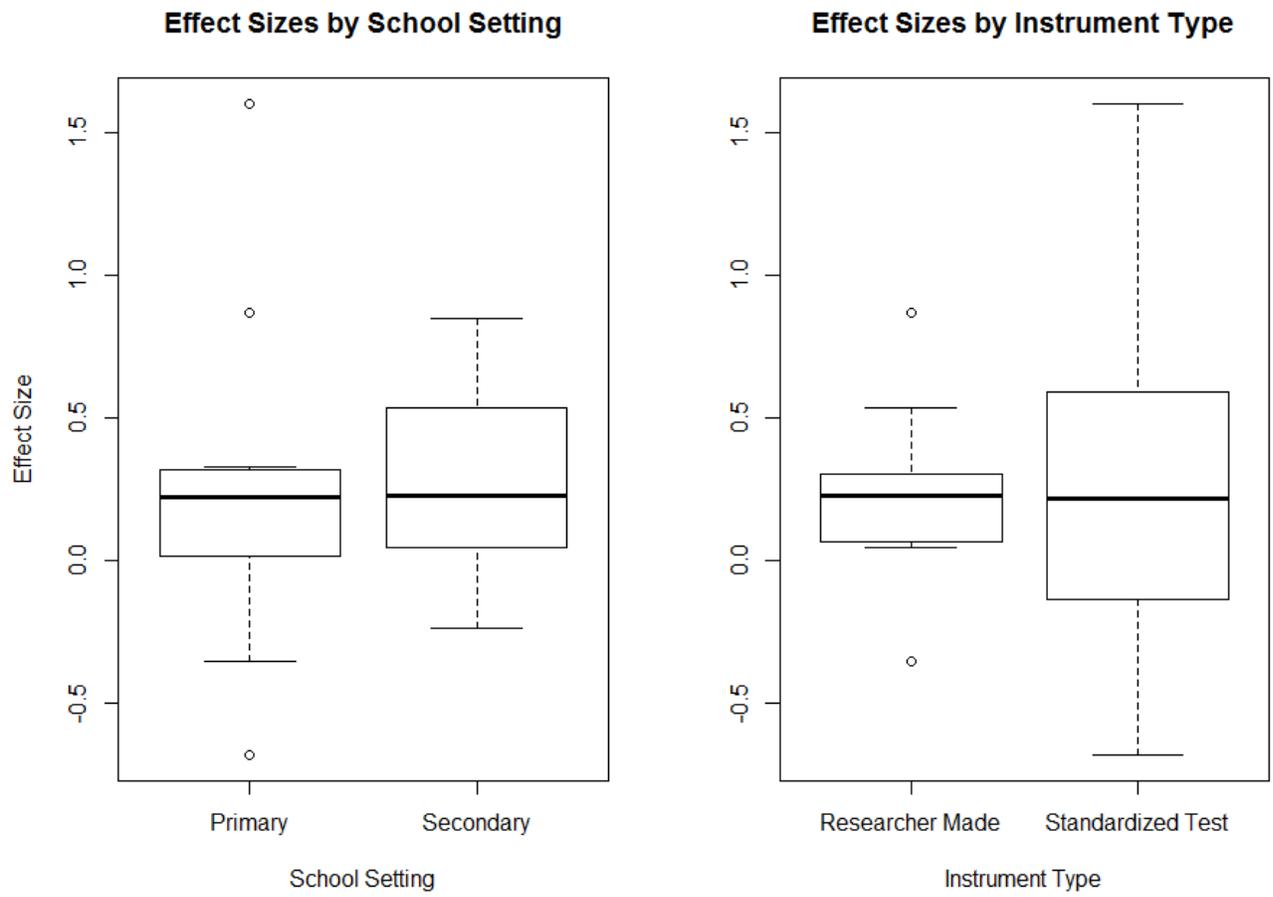
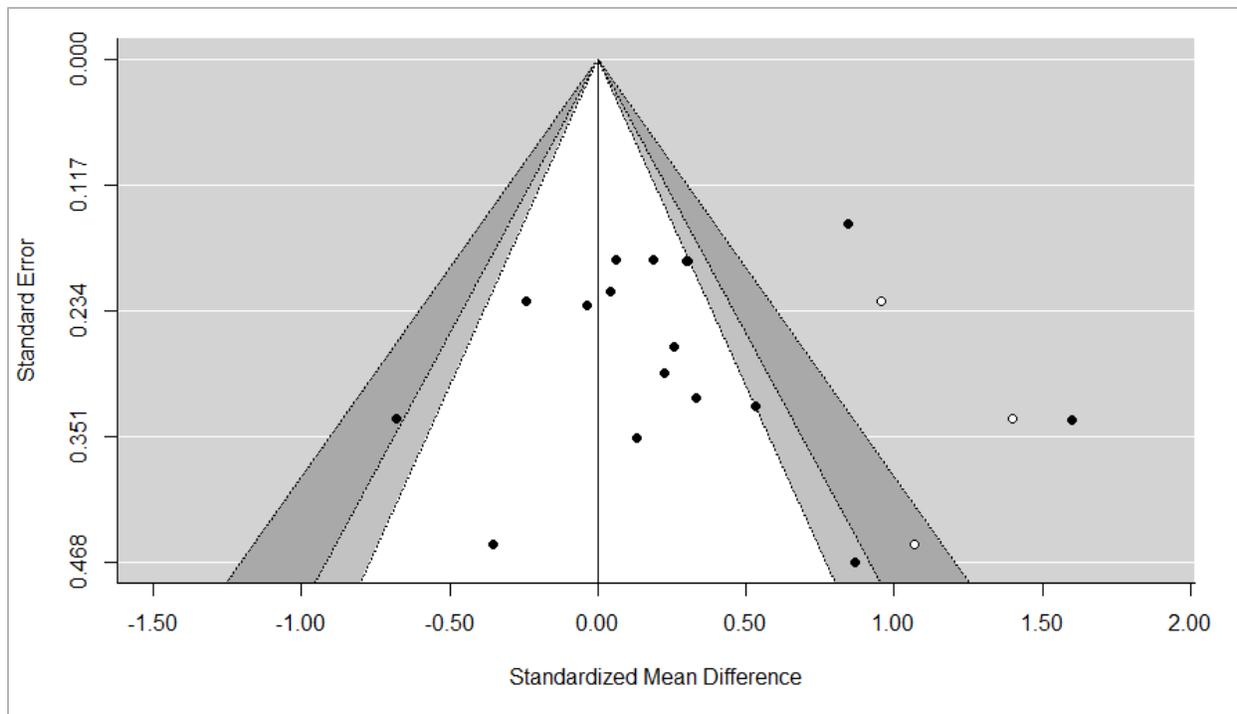


Figure 2. Boxplots for each group of both moderators used in analyses.



*Figure 3.* Funnel plot of effect sizes with approximate 90% (white), 95% (light gray), and 99% (dark gray) confidence intervals. Filled circles are collected effects and unfilled circles represent imputed effects based on the Trim and Fill method (Duval & Tweedie, 2000).

Table 1

*Study Characteristics included in the Meta-Analysis*

<i>Study</i>	<i>Sample Size</i>	<i>School Setting</i>	<i>Country</i>	<i>Instrument</i>	<i>Game(s)</i>	<i>Intervention Duration (minutes)</i>	<i>Teachers preparation for game-based instruction (minutes)</i>
Din & Caleo (2000)	47	Primary	USA	Standard academic test	Lightspan, Sony Play Station (~ 40 different games)	2200	NI
Hawkins (2008)	103	Primary	USA	Standard academic test	MySims Wii, Nintendo Wii	300	120
Ke & Grabowski (2007)	125	Primary	USA	Standard academic test	ASTRA EAGLE (a series of web-based computer games; academic content is based on the Pennsylvania System of School Assessment (PSSA) standards for mathematics)	320	120
Kebritchi, Hirumi, & Bai (2010)	193	Secondary	USA	Other	DimensionM	540	60
Pareto, Haake, Lindstrom, Sjoden & Gulz (2012)	38	Primary	Sweden	Large Scale/Standardized Survey /questionnaire	Teachable Agents game (researcher-made game)	2205	NI
Sedig (2008)	79	Secondary	Canada	Research specific test	Super Tangrams (researcher-made game)	400	60

Sung, Chang & Lee (2008)	60	Primary	Taiwan	Research specific test	<i>SoRT</i> (researcher-made game) <i>My First Math Adventure – Counting and Classification</i> (Fu, 1999) <i>Toby’s IQ Training Camp – the Seed of Logic</i> (SMEC, 1999).	60	NI
Van Eck & Dempsey (2002)	112	Secondary	USA	Research specific test	A simulation game (researcher-made game; academic content is based on the National Council of Teachers of Mathematics standards)	50	NI
Ke (2008)	358	Primary	USA	Standard academic test	ASTRA EAGLE (a series of web-based computer games; academic content is based on the Pennsylvania System of School Assessment (PSSA) standards for mathematics)	360	60
Delacruz (2010)	103	Primary	USA	Large Scale/Standardized Survey /questionnaire	<i>Save Patch</i> (developed by the National Center for Evaluation, Standards, and Student Testing (CRESST) and students in a game program at the University of Southern California's (USC) Game Innovation Lab)	25 to 40	NI
Gelman (2010)	80	Secondary	USA	Research specific test	Brain Age 2, Nintendo DS	675	180

Ke (2006)	487	Primary	USA	Research specific test	ASTRA EAGLE math games (a series of web-based computer games; academic content is based on the Pennsylvania System of School Assessment (PSSA) standards for mathematics)	360	60
King (2011)	128	Secondary	USA	NI	VmathLive (academic content is based on the National Council of Teachers of Mathematics (NCTM) standards)	NI	180

Note. *NI* = no information