WWC Single Study Review

A review of the design and summary of findings for an individual study

ies INSTITUTE OF EDUCATION SCIENCES

November 2015

WWC Review of the Report "The Effects of Math Video Games on Learning"[1]

The findings from this review do not reflect the full body of research evidence on fractions-related math video games.

What is this study about?

The study authors examined the effectiveness of four fractions-related math video games in improving students' fractions knowledge. The four games were titled *Wiki Jones*, *Save Patch*, *Tlaloc's Book*, and *Rosie's Rates*. The authors compared these fractions-knowledge-focused games to four different video games that emphasized equation solving. Both sets of video games were developed by the Center for Advanced Technology in Schools (CATS), an IES-funded center focused on developing and evaluating computer games to improve students' pre-algebra knowledge.

The study is a cluster randomized controlled trial. In response to an author query, the authors indicated that 80 sixth-grade classrooms within 26 schools across nine school districts were originally randomly assigned to condition: 42 to the intervention group and 38 to the comparison group.[2] Each classroom was led by a different teacher. Prior to the intervention, teachers in both groups received 3 hours of professional development on the games and how to connect them to their instruction.

The analytic sample included 59 classrooms (30 in the intervention and 29 in the comparison group) in 23 schools. In response to an author query, the authors indicated that 1,571 students were enrolled in these participating classrooms at random assignment; 1,468 of these students were included in the analytic sample (759 in the intervention group and 709 in the comparison group).

The study examined the impact of fractions-related video games using a researcher-designed measure of fractions knowledge. This assessment was administered 1 week following the last day that classrooms completed the final video game.[3]

Features of the Center for Advanced Technology in Schools (CATS) fractions-related math video games

The intervention condition consisted of four video games—*Wiki Jones*, *Save Patch*, *Tlaloc's Block*, and *Rosie's Rates*—designed by CATS for use in this study as part of an initiative to create video games to improve students' knowledge of pre-algebra concepts. These four video games focused on fractions concepts including: whole units, numerator and denominator, understanding the number line, fractions computation, and slope.

Students played the games independently in a predetermined sequence. Students played the four games over ten 40-minute classroom periods. Three games were played for two periods each. The fourth game, *Save Patch*, was played for four periods. Students completed an immediate game-specific assessment after each game ended, which was not an outcome examined in the study. Prior to the intervention, teachers in both groups received 3 hours of professional development on the games themselves and the associated concepts, as well as how to integrate the games with their curriculum.

What did the study find?

The study authors reported, and the WWC confirmed, statistically significant positive effects of the CATS fractions-related math video games on fractions knowledge in sixth-grade students.

WWC Rating

The research described in this report meets WWC group design standards without reservations

This study is a cluster randomized controlled trial with low levels of sample attrition.

Appendix A: Study details

Chung, G. K. W. K., Choi, K., Baker, E., & Cai, L. (2014). *The effects of math video games on learning: A randomized evaluation study with innovative impact estimation techniques*. National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles. http://files.eric.ed.gov/fulltext/ED555700.pdf.

Setting

The study took place in 26 schools across nine school districts in unspecified locations.

Study sample

The initial sample included 80 sixth-grade classrooms/teachers in 26 schools that were randomly assigned to the intervention and comparison conditions. In total, 42 classrooms were initially assigned to the intervention group, and 38 were assigned to the comparison group. The final analytic sample, after attrition, included 59 classrooms (30 in the intervention group and 29 in the comparison group) from 23 schools. From a total of 1,571 students who were enrolled in these participating classrooms at random assignment (808 in the intervention group and 763 in the comparison group), the analytic sample includes data for 1,468 students (759 in the intervention group and 709 in the comparison group). Half of students in participating classrooms were female. Nearly half (49%) of the students were Hispanic, 24% were White, 5% were African American, 4% were Asian or Pacific Islander, 2% American Indian or Alaskan Native, and 5% reported "other" race or ethnicity.

Intervention group

The fraction video games intervention included four video games (entitled *Wiki Jones*, *Save Patch*, *Tlaloc's Book*, and *Rosie's Rates*) developed by the Center for Advanced Technology in Schools (CATS) as part of an initiative to create video games to improve students' knowledge of pre-algebra concepts. The games were designed to teach rational numbers and fractions in a predetermined order over 10 days. The games were designed with the intention of keeping students engaged and on task as they learned key fraction-related concepts such as: whole units, numerator, denominator, identifying fractions using a number line, adding fractions, multiplicative inverse operations, and slope. Students played the games independently. Students used three of the four games over two 40-minute classroom periods, and one game *(Save Patch)* over four classroom periods. During these classroom periods, students played the game and took an immediate, game-specific posttest upon completion of each game.

Comparison group

The comparison condition was similar to the intervention condition, except the four video games (entitled *Monster Line*, *Expresso*, *Zooples in Space*, and *AlgebRock*) focused on equation solving. The games for the comparison condition were also developed by CATS and used in a predetermined order over 10 days. These games were designed to keep students engaged and on task as they learned key concepts related to solving equations such as: operations on positive and negative integers, manipulating and transforming math expressions, equality and additive inverse operations, and solving single and two-step equations. Students played the games independently. Students used three of the four games over two 40-minute classroom periods, and one game (*Zooples in Space*) over four classroom periods. During these classroom periods, students played the game and took an immediate, game-specific posttest upon completion of each of the games.

Outcomes and measurement

Students completed a general fractions knowledge assessment 1 week following the end of the last game. This assessment was designed by researchers for use in this study. For a more detailed description of this outcome measure, see Appendix B.

Support for implementation

Before students played the games in the intervention, teachers in both the intervention and comparison groups took part in 3 hours of professional development. Teachers received training associated with the specific games in the condition to which they were randomly assigned. During the professional development time, teachers were first given instruction on the concepts and games; then the teachers played the video games; and finally, teachers discussed ways to connect the games to their curriculum.

Reason for review

This study was identified for review by the WWC because it was supported by a grant to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) (Principal Investigator: Eva Baker) from the National Center for Education Research (NCER) at the Institute of Education Sciences (IES).

Appendix B: Outcome measures for the mathematics achievement domain

Mathematics achievement	
Fractions knowledge	The measure was developed for use in this study by a team that included the lead study author and was intended to measure the same knowledge emphasized by the games (Vendlinski et al., 2010). The measure focuses on two key ideas related to the addition of rational numbers drawn from the literature and discussions with math educators: (1) "only identical units can be added to create a single numerical sum," and (2) "the size of a rational number is relative to how one whole unit is defined." According to the authors, the measure contains items comparable to other standardized assessments. There were 22 and 23 items on the pre- and posttest, respectively. The outcome has test-retest reliability of 0.90. The authors also reported classical reliability estimates of 0.80.

Table Notes: The study authors calculated this outcome in two ways. The WWC review focused on the results from the classical measurement approach that used summed scores across test items. The authors also created an alternative version of this outcome using a multi-level two-tier item factor model to account for differences based on group assignment and variation across schools. The WWC did not report on the impact analysis results based on this latter method because the measure implicitly measures gain scores, which lead to effect sizes that are not comparable to effect sizes the WWC reports in similar studies. Additional data on the fractions knowledge outcome measure were collected from the following source: Vendlinski, T. P., Delacruz, G. C., Buschang, R. E., Chung, G. K. W. K., & Baker, E. (2010). *Developing high-quality assessments that align with instructional video games*. National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles. http://files.eric.ed.gov/fulltext/ED512655.pdf.

Appendix C: Study findings for the mathematics achievement domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mathematics achievement								
Fractions knowledge	sixth grade students	23 schools/ 1,468 students	10.69 (5.22)	9.57 (4.59)	1.12	0.23	+9	.001
Domain average for mathematics achievement						**0.23**	**+9**	**Statistically significant**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding.

Study Notes: The WWC did not need to make corrections for clustering, multiple comparisons, or to adjust for baseline differences. The *p*-value presented here was reported in the original study. The study reports *p*-values and impact estimates from this multi-level model. The effect size was calculated by the WWC based on the study-reported impact estimate and unadjusted standard deviations supplied by the authors. The WWC calculated the intervention group mean by adding the study-reported impact estimate of the intervention to the comparison group posttest mean, which the authors provided upon request. This study is characterized as having a statistically significant positive effect because the effect for the one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant. For more information, please refer to the WWC Standards and Procedures Handbook (version 3.0), p. 26.

Endnotes

[1] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the authors) to assess whether the study design meets WWC group design standards. The review reports the WWC's assessment of whether the study meets WWC group design standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the Primary Mathematics review protocol (version 3.1).

[2] In response to an author query, the authors reported that the baseline sample size of 62 classrooms reported in the study corresponds to the number of classrooms that started the intervention—not the number assigned to conditions. In the analytic sample, a total of 14 of the schools included both intervention and comparison classrooms, six schools had only intervention classrooms, and three schools had only comparison classrooms. In some schools, only one classroom was assigned to condition. In response to an author query, the authors reported that this happened for a variety of reasons; for example, because only one teacher in the school agreed to participate, or because a teacher did not attend the required professional development.

[3] The study authors measured this outcome in two ways. This WWC review focused on the results from the classical measurement approach that focused on the sum of scores across items. The authors also created an alternative version of this outcome using the same data and multilevel item response theory, which is not a focus of this WWC review. See the table notes in Appendix B for more information.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2015, November). *WWC review of the report: The effects of math video games on learning: A randomized evaluation study with innovative impact estimation techniques.* Retrieved from http://whatworks.ed.gov

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analytic sample groups are similar on observed characteristics defined in the review area protocol.

Improvement index Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which study participants are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which eligible study participants are randomly assigned to intervention and comparison groups.

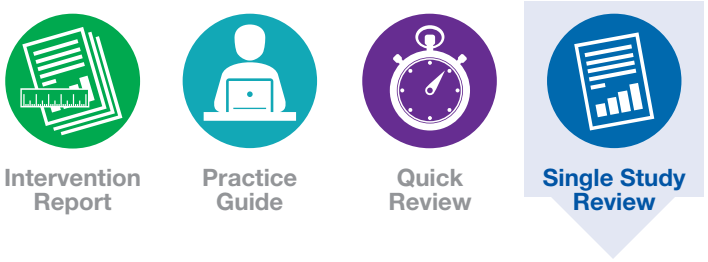
Single-case design (SCD) A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample are spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < .05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 3.0) for additional details.



A **single study review** of an individual study includes the WWC's assessment of the quality of the research design and technical details about the study's design and findings.

This single study review was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.