WWC Intervention Report

A summary of findings from a systematic review of the evidence

ies INSTITUTE OF EDUCATION SCIENCES

Primary Mathematics August 2017*

I CAN Learn®

Report Contents

This intervention report presents findings from a systematic review of *I CAN Learn*® conducted using the WWC Procedures and Standards Handbook (version 3.0) and the Primary Mathematics review protocol (version 3.1).

Intervention Description[1]

I CAN Learn® is a computer-based math curriculum for students in middle school, high school, and college. It provides math instruction through a series of interactive lessons that students work on individually at their own computers. Students move at their own pace and must demonstrate mastery of each concept before progressing to the next one. Classroom teachers may provide individual, small-group, or whole-class instruction based on students' performance on the software program. This review focuses on studies of *I CAN Learn*®'s primary courses, which include *Fundamentals of Math* and *Pre-Algebra*.

Research[2]

The What Works Clearinghouse (WWC) identified one study of *I CAN Learn*® that both falls within the scope of the Primary Mathematics topic area and meets WWC group design standards. No studies meet WWC group design standards without reservations, and one study meets WWC group design standards with reservations. This study included 9,886 students in eighth grade in one school district.

The WWC considers the extent of evidence for *I CAN Learn*® on the mathematics test scores of students in eighth grade to be small for the mathematics achievement domain, the only domain examined for studies reviewed under the Primary Mathematics topic area.[3] (See the Effectiveness Summary on p. 4 for more details.)

Effectiveness

I CAN Learn® had no discernible effects on the mathematics test scores of eighth-grade students in primary mathematics courses.

Table 1. Summary of findings[4]

Outcome domain	Rating of effectiveness	Improvement index (percentile points) Average	Improvement index (percentile points) Range	Number of studies	Number of students	Extent of evidence
Mathematics achievement	No discernible effects	+2	na	1	9,886	Small

na = not applicable

Intervention Information

Background

JRL Enterprises, Inc. is the developer and distributor of *I CAN Learn*®. Address: 912 Constantinople Street, New Orleans, LA 70115. Email: info@icanlearn.com. Website: http://www.icanlearn.com. Telephone: (504) 263-1380.

Intervention details

I CAN Learn® is a full-curriculum mathematics software program. It is available in online learning models and in traditional classroom models. *I CAN Learn*® provides math instruction through a series of computer-based interactive lessons that students work on individually at their own computers in middle school, high school, and college classrooms. Students typically start with a placement test, which is used to assess their current level of knowledge and customize their lesson plan. Each lesson includes a warm-up activity, lesson presentation, guided practice, and a quiz to ensure mastery of the lesson content. A virtual teacher in the software presents the lessons and demonstrates how to solve problems. Students then work at their own pace and must demonstrate mastery of each concept before progressing to the next one. Classroom teachers may provide individual, small-group, or whole-class instruction based on students' performance on the software program. Printed textbooks are available to supplement the electronic courseware.

I CAN Learn® courses include *Fundamentals of Math*, *Pre-Algebra, Algebra*, and *Geometry*. *I CAN Learn*® *Fundamentals of Math* and *Pre-Algebra* are examined as part of this primary mathematics review. According to the developer, all four courses align to the National Council of Teachers of Mathematics standards and can be customized to meet state- or district-specific standards. Each course contains 20 to 180 self-paced, mastery-based lessons. Lessons include multimedia content, such as instructional videos and animation. In addition to the instructional content, *I CAN Learn*® enables teachers to conduct classroom administration tasks through the *I CAN Learn*® *Classroom Explorer Class Management System*, which tracks students' attendance, homework, and test grades, and can create an individualized learning path for each student. In addition, teachers can use the system to generate a variety of reports that support instructional decisions, such as identifying areas in which students need assistance; grouping students; and tracking attendance, homework, and grades.

Cost

The cost of *I CAN Learn*® depends on its configuration and terms of support. As of March 2017, using a school's existing hardware, subscriptions provide access to more than 600 lessons and cost from $10 per student for site licenses to $53.48 for an individual student license. Varying support plans, including training, professional development, curriculum alignments, implementation planning, and other pedagogical support, are available and cost from $400 to $20,000 per year. More detailed information about cost is available from the developer.

Research Summary

The WWC identified seven eligible studies that investigated the effects of *I CAN Learn*® on the mathematics achievement of primary students. An additional 33 studies were identified but do not meet WWC eligibility criteria (see the Glossary of Terms in this document for a definition of this term and other commonly used research terms) for review in this topic area. Citations for all 40 studies are in the References section, which begins on p. 5.

Table 2. Scope of reviewed research

Grade	8
Delivery method	Whole class
Intervention type	Curriculum

The WWC reviewed seven eligible studies against group design standards. None of the seven studies is a randomized controlled trial that meets WWC group design standards without reservations. One study uses a quasi-experimental design that meets WWC group design standards with reservations. This report summarizes the one study. The remaining six studies do not meet WWC group design standards.

Summary of studies meeting WWC group design standards without reservations

No studies of *I CAN Learn*® met WWC group design standards without reservations.

Summary of study meeting WWC group design standards with reservations

Kerstyn (2002) used a quasi-experimental design to examine the effects of *I CAN Learn*® on eighth-grade students in 36 middle schools in one school district in the 2000–01 and 2001–02 school years. Twelve of the district's 36 middle schools implemented *I CAN Learn*® *Pre-Algebra* and *I CAN Learn*® *Algebra* in 2000–01. During the 2001–02 school year, two more middle schools in the district implemented the program. The district staff matched the *I CAN Learn*® classrooms to similar classrooms that used their district's traditional curriculum. Each year of the study, the author presented the results separately for two types of classes: pre-algebra (called MJ-3) and advanced pre-algebra (called MJ-3 Advanced).[5] Across both school years and types of classes, there were a total of 524 classrooms and 9,886 students in the study: 143 classrooms and 2,349 students in the intervention group and 381 classrooms and 7,537 students in the comparison group. The study used the Florida Comprehensive Assessment Test math assessment to measure eighth-grade mathematics achievement outcomes.[6] The study did not specify the edition of *I CAN Learn*® used.

Effectiveness Summary

The WWC review of *I CAN Learn*® for the Primary Mathematics topic area includes student outcomes in one domain: mathematics achievement. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *I CAN Learn*® on primary students. Additional comparisons are presented as supplemental findings in Appendix D. The supplemental findings do not factor into the intervention's rating of effectiveness. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 17.

Summary of effectiveness for the mathematics achievement domain

Table 3. Rating of effectiveness and extent of evidence for the mathematics achievement domain

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *mathematics achievement* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 9,886 students in 36 schools reported evidence of effectiveness in the *mathematics achievement* domain.

One study that met WWC group design standards with reservations reported findings in the mathematics achievement domain.

Kerstyn (2002) reported, and the WWC confirmed, no statistically significant difference between *I CAN Learn*® and the comparison group on the mathematics achievement domain for three analytic samples: MJ-3 and MJ-3 Advanced classrooms in the 2000–01 school year and the MJ-3 Advanced students in the 2001–02 school year. For the MJ-3 students in the 2001–02 school year, the author reported a positive and statistically significant difference between the *I CAN Learn*® group and the comparison group in the mathematics achievement domain. However, after adjusting this finding for multiple comparisons (that is, changing significance levels to take into account several comparisons), the WWC found that the difference was no longer statistically significant. None of the comparisons had an effect size that was large enough to be considered substantively important according to WWC criteria (that is, an effect size of at least 0.25). The WWC characterizes this study finding as an indeterminate effect.

Thus, for the mathematics achievement domain, one study of *I CAN Learn*® showed an indeterminate effect. This results in a rating of no discernible effects, with a small extent of evidence.

References

Studies that meet WWC group design standards without reservations

None.

Study that meets WWC group design standards with reservations

Kerstyn, C. (2002). *Evaluation of the I CAN Learn mathematics classroom: Second year of implementation (2001–2002 school year).* Tampa, FL: Hillsborough County Public Schools

Additional sources:

Kerstyn, C. (2001). *Evaluation of the I CAN Learn mathematics classroom: First year of implementation (2000–2001 school year)*. Tampa, FL: Hillsborough County Public Schools.

Kerstyn, C. (2004). *Teachers' mathematics preparation and eighth grade student mathematics achievement: Can an integrated learning system provide support when teachers' professional preparation is limited?* (Doctoral dissertation, University of South Florida).

Studies that do not meet WWC group design standards

Barrow, L., Markman, L., & Rouse, C. E. (2009). Technology's edge: The educational benefits of computer-aided instruction. *American Economic Journal: Economic Policy, 1*(23), 52–74. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Additional sources:

Barrow, L., Markham, L., & Rouse, C. E. (2007). *Technology's edge: The educational benefits of computer-aided instruction* (Working Paper Series: WP-07-17). Chicago, IL: Federal Reserve Bank of Chicago.

Barrow, L., Markman, L., & Rouse, C. E. (2008). *Technology's edge: The educational benefits of computer-aided instruction* (NBER Working Paper 14240). Cambridge, MA: National Bureau of Economic Research.

DeLoach, R. M. (2011). *A summative evaluation of the effectiveness of classroom-embedded, individualistic, computer-based learning for middle school students placed at academic risk in schools with high proportion of Title I eligible students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3453582) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Gill, J. C., & Gifford, C. S. (2001). *Evaluation of Jefferson Parish technology grant: I CAN Learn Algebra I.* Unpublished manuscript, University of New Orleans, LA. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Additional sources:

Brooks, C. (1999). *Evaluation of Jefferson Parish technology grant: I CAN Learn Algebra I. Jefferson Parish Public Schools.* New Orleans, LA: University of New Orleans, Department of Educational Leadership.

Brooks, C. (2000). *Evaluation of Jefferson Parish technology grant: I CAN learn Algebra I.* New Orleans, LA: University of New Orleans, Department of Educational Leadership.

JRL Enterprises, Inc. (n.d.). *I CAN Learn® success in Louisiana: I CAN Learn® students outperform traditionally-taught students.* New Orleans, LA: Author.

I CAN Learn. (n.d.). *I CAN Learn® vs traditionally-taught students: Fort Worth, Texas 2002.* New Orleans, LA: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Additional source:

Kirby, P. C. (2005). *Comparison of I CAN Learn and traditionally-taught 7th and 9th grade student performance on the Texas criterion-referenced tests, 2000–2004.* New Orleans, LA: I CAN Learn.

Kirby, P. C. (2004a). *Comparison of I CAN Learn® and traditionally-taught 8th grade general math student performance on the California Standards Test, Spring 2004.* New Orleans, LA: I CAN Learn. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Kirby, P. C. (2004b). *Texas district performance on TAAS and TAKS, 1999–2003. I CAN Learn® in Fort Worth Independent School District.* New Orleans, LA: Ed-Cet, Inc. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

Studies that are ineligible for review using the Primary Mathematics Evidence Review Protocol

Aql, M. M. (2011). Effects of computer integrated instruction on student achievement in eighth grade mathematics. (Doctoral dissertation, Baker University, Baldwin City, KS). The study is ineligible for review because it is out of the scope of the protocol.

Buckler, M. L. (2007). *Comparison study of 8th grade math MAP scores of four Missouri middle schools using the I CAN Learn® math lab in academic years 2005 and 2006*. Unpublished master's thesis, Northwest Missouri State University, Maryville. The study is ineligible for review because it is out of the scope of the protocol.

Center for Data-Driven Reform in Education. (2009). *Effective programs in middle and high school mathematics: A best-evidence synthesis [Educator's summary*]. Baltimore, MD: Author. The study is ineligible for review because it does not use an eligible design.

Cheung, C. K., & Slavin, R. E. (2011). *The effectiveness of educational technology applications for enhancing mathematics achievement in K–12 classrooms: A meta-analysis.* Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education. Retrieved from https://eric.ed.gov/?id=ED527573 The study is ineligible for review because it does not use an eligible design.

I CAN Learn. (n.d.). *California report for teachers: 2008 CST for general mathematics.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

I CAN Learn. (n.d.). *Comparisons between Orleans Parish and statewide 8th grade LEAP passing rates.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

I CAN Learn. (n.d.). *I CAN Learn success in Birmingham.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

I CAN Learn. (n.d.). *Results in Los Angeles Unified School District: Sepulveda Middle School.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

I CAN Learn. (2006). *I CAN Learn math in Missouri: 8th grade.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

I CAN Learn. (2007a). *New York City Middle School 385.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

I CAN Learn. (2007b). *Northridge Middle School Algebra Readiness students significantly outperform traditionally taught students.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

JRL Enterprises, Inc. (n.d.a). *California: Los Angeles Unified School District.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

JRL Enterprises, Inc. (n.d.b). *I CAN Learn® success in California: I CAN Learn® students outscore traditionally-taught students.* New Orleans, LA: Author. The study is ineligible for review because it is out of the scope of the protocol.

Additional source:

Oescher, J. (2002). *I CAN Learn® education success in California*. New Orleans, LA: Author.

JRL Enterprises, Inc. (n.d.c). *Oklahoma schools see big gains on OCCT scores using the I CAN Learn program.* New Orleans, LA: Author. The study is ineligible for review because it does not use an eligible design.

Kirby, P. C. (n.d.). *I CAN Learn results in Mississippi.* New Orleans, LA: Ed-Cet, Inc. The study is ineligible for review because it is out of the scope of the protocol.

Kirby, P. C. (2003). *I CAN Learn® results for Orleans parish schools 2001–2002.* New Orleans, LA: I CAN Learn. The study is ineligible for review because it is out of the scope of the protocol.

Kirby, P. C. (2004c). *California Hayward School District, Bret Harte Middle School.* New Orleans, LA: I CAN Learn. The study is ineligible for review because it does not use an eligible design.

Kirby, P. C. (2004d). *Comparison of I CAN Learn® and traditionally-taught 8th grade student performance on the Georgia criterion-referenced competency test.* New Orleans, LA: I CAN Learn. The study is ineligible for review because it is out of the scope of the protocol.

Kirby, P. C. (2004e). *I CAN Learn® in Collier County, FL.* New Orleans, LA: Ed-Cet, Inc. The study is ineligible for review because it is out of the scope of the protocol.

Kirby, P. C. (2005). *I CAN Learn® results in Milwaukee, Wisconsin.* New Orleans, LA: Ed-Cet, Inc. The study is ineligible for review because it is out of the scope of the protocol.

Kirby, P. C. (2006a). *I CAN Learn® Algebra I in Catoosa County, Georgia: Final report.* New Orleans, LA: I CAN Learn. The study is ineligible for review because it is out of the scope of the protocol.

Additional source:

Kirby, P. C. (2005). *I CAN Learn® Algebra I in Catoosa County, Georgia*. New Orleans, LA: I CAN Learn.

Kirby, P. C. (2006b). *I CAN Learn® in Orleans parish public schools: Effects on LEAP 8th grade math achievement, 2003–2004.* New Orleans, LA: I CAN Learn. This study is ineligible for review because it is out of the scope of the protocol.

Additional source:

Kirby, P. C. (2004). *I CAN Learn® in Orleans parish public schools: Effects on LEAP 8th grade math achievement, 2003–2004.* New Orleans, LA: I CAN Learn.

Kirby, P. C. (2006c). *I CAN Learn® in Orleans Parish Public Schools effects on LEAP 10th grade math achievement, 2003–2004.* New Orleans, LA: Ed-Cet, Inc. This study is ineligible for review because it is out of the scope of the protocol.

Kirby, P. C. (2006d). *I CAN Learn® Algebra I in Central Falls, Rhode Island, 2005–2006.* New Orleans, LA: Ed-Cet, Inc. The study is ineligible for review because it is out of the scope of the protocol.

Oescher, J., & Kirby, P. C. (2004). *I CAN Learn results in Dallas, Texas 9th grade 2003–2004.* New Orleans, LA: I CAN Learn. The study is ineligible for review because it is out of the scope of the protocol.

Parrott, D. E. (2005). *A study of comparison for students in a computerized mathematics classroom I CAN Learn® and students who are in a traditionally taught classroom.* Unpublished master's thesis, Northwest Missouri State University, Maryville. The study is ineligible for review because it is out of the scope of the protocol.

Saultz, A. (2012). *Programs for middle school math: An inventory of existing technology.* Stanford, CA: Policy Analysis for California Education. Retrieved from https://eric.ed.gov/?id=ED532518 The study is ineligible for review because it does not use an eligible design.

Scafide, K. (2004). *Effects of I CAN Learn® on math achievement in Gwinnett County Middle School.* New Orleans, LA: I CAN Learn. The study is ineligible for review because it is out of the scope of the protocol.

Slavin, R. E., Lake, C., & Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research, 79*(2), 839–911. The study is ineligible for review because it does not use an eligible design.

Additional source:

Slavin, R. E., Lake, C., & Groff, C. (2008). *Effective programs in middle and high school mathematics: A best-evidence synthesis.* Baltimore, MD: Johns Hopkins University, Center for Data-Driven Reform in Education. Retrieved from https://eric.ed.gov/?id=ED527635

Stokes, S. D. (2011). *A multi-metric assessment on the impact of I Can Learn® (ICL) multimedia on actual and perceived student achievement in developmental mathematics* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3454223) The study is ineligible for review because it does not use a sample aligned with the protocol.

Williams, K. C. (2015). *The effects of computer-assisted instruction on the mathematics achievement of students with emotional and behavioral disorders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3704169) The study is ineligible for review because it does not use a sample aligned with the protocol.

Woods, C. S. (2012). *Effects of I Can Learn (ICL) on the attitudes and academic achievement of eighth grade math students in two rural Tennessee school districts.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3519082) The study is ineligible for review because it is out of the scope of the protocol.

Zinn, L. F., & Bolton, J. W. (2006). *EETT grant evaluation: Year 2 report. Williamsport area school district.* New Orleans, LA: I CAN Learn. The study is ineligible for review because it is out of the scope of the protocol.

Appendix A: Research details for Kerstyn (2002)

Kerstyn, C. (2002). *Evaluation of the I CAN Learn mathematics classroom: Second year of implementation (2001–2002 school year)*. Tampa, FL: Hillsborough County Public Schools.

Additional sources:

Kerstyn, C. (2001). *Evaluation of the I CAN Learn mathematics classroom: First year of implementation (2000–2001 school year)*. Tampa, FL: Hillsborough County Public Schools.

Kerstyn, C. (2004). *Teachers' mathematics preparation and eighth grade student mathematics achievement: Can an integrated learning system provide support when teachers' professional preparation is limited?* (Doctoral dissertation, University of South Florida).

Table A. Summary of findings

Meets WWC group design standards with reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Mathematics achievement	524 classrooms/ 9,886 students	+2	No

Setting

The study was conducted in 524 classrooms across 36 schools in the Hillsborough County Public School District in Florida. The study occurred in the 2000–01 and 2001–02 school years.

Study sample

The district implemented *I CAN Learn® Algebra* and *I CAN Learn® Pre-Algebra* in select classrooms within 12 middle schools during the 2000–01 school year. In the 2001–02 school year, the 12 middle schools continued implementing *I CAN Learn®*, and two more middle schools began implementing it. For the evaluation, within each study school, district staff selected comparison classrooms that used traditional instruction and matched them on several factors including: students' prior achievement, time of day, instructional time, class size, and proportion of minority students. This review is based on the *I CAN Learn® Pre-Algebra* classrooms in the first (2000–01) and second year (2001–02) of the study, which included students in all of the district's 36 middle schools in MJ-3 and MJ-3 Advanced classrooms with pre- and posttest data. In 2000–01, the MJ-3 analysis included 32 *I CAN Learn®* classrooms and 32 comparison classrooms. The MJ-3 Advanced analysis included 10 *I CAN Learn®* classrooms and 10 comparison classrooms. In 2001–02 school year, the MJ-3 analysis included 64 *I CAN Learn®* classrooms and 264 comparison classrooms. The MJ-3 Advanced analysis included 37 *I CAN Learn®* classrooms and 75 comparison classrooms.

There were 9,886 students in the study: In the 2000–01 school year, there were 1,420 MJ-3 students and 430 MJ-3 Advanced students. Approximately 50% of the students were male, and about 45% were categorized as qualifying for free or reduced-price meals. In terms of racial/ethnic composition, the intervention group was 37% White, 33% Hispanic, and 30% Black, compared to the comparison group, which was 47% White, 28% Hispanic, and 25% Black. In the 2001–02 school year, there were 5,957 MJ-3 students and 2,079 MJ-3 Advanced students. Less than 10% of the students were categorized as Exceptional Education (ESE), Limited English Proficiency (LEP), or home-schooled. Approximately 50% of comparison

students were male. About 65% of intervention students and 39% of comparison students were categorized as qualifying for free or reduced-price meals. In terms of racial/ethnic composition, the intervention group was 34% White, 34% Hispanic, and 28% Black, compared to the comparison group, which was 53% White, 22% Hispanic, and 20% Black.

The author conducted analyses using alternate samples in the second year of the study (2001–02), and the additional second year analytic samples that met standards are also presented as supplemental findings in Appendix D. These include an alternate sample of students in all 36 of the middle schools in the district and a sample of students in the 14 middle schools that piloted *I CAN Learn®*.[7] The supplemental findings do not factor into the intervention's rating of effectiveness.

Intervention group

Intervention students were taught using *I CAN Learn®* as the primary source of math instruction for the entire academic school year. The curriculum includes 109 lessons, each of which has a five-part format that includes a warm-up activity, lesson presentation, journal activity, guided practice, and a quiz to ensure mastery of the lesson content. Students are expected to complete the lessons individually and at their own pace using interactive software with a virtual teacher that presents the multimodal lessons and demonstrates how to solve a problem if students make errors. In each class, a classroom teacher supported students with the lessons. The study did not specify which edition of the curriculum was used.

Comparison group

Comparison students used a traditional math curriculum already in place in the district. The author did not describe or name the comparison curriculum.

Outcomes and measurement

The primary outcome measure is the Florida Comprehensive Assessment Test (FCAT) math exam aligned with Florida's Sunshine State Standards (SSS). The outcome is measured using the eighth-grade scale score. The seventh-grade FCAT Norm Reference Test-Norm Curve Equivalent (NRT-NCE) scores were used as the baseline assessment for students in the analytic sample.

The study also presented outcomes on the district's first semester Cumulative Test in the 2001–02 school year, which students take at the end of the first semester. Because this assessment measures student achievement mid-year, it is considered a supplemental outcome that does not factor into the intervention's rating of effectiveness and is presented in Appendix D. For a more detailed description of these eligible outcome measures, see Appendix B.

The study presents several outcomes that are ineligible for review under the Primary Math review protocol, including teacher attitudes, parent attitudes, and student attitudes.

Support for implementation

The study did not specify how much training intervention teachers received. The district contracted with JRL Enterprises, Inc. to use *I CAN Learn®* during the year prior to the study (1999–2000). To implement the curriculum, classrooms were equipped with desks, computer equipment, and electrical connectivity. JRL Enterprises, Inc. provided maintenance on the equipment and technical and instructional support to teachers.

Appendix B: Outcome measures for the mathematics achievement domain

Mathematics achievement	
Cumulative Test	The Cumulative Test is a pre-algebra assessment developed by the district. All students take the Cumulative Test at the end of the each semester, but the school determines how the test score is used in the students' overall assessment (as cited in Kerstyn, 2001). In response to a WWC query, the author indicated that the exam has an internal consistency between .84 and .88. This outcome is only reported as a supplemental finding in Appendix D.
Florida Comprehensive Assessment Test (FCAT) Mathematics	The FCAT Mathematics exam is a standardized assessment that includes items aligned to all five content strands in the Florida Sunshine State Standards: (1) number sense, concepts, and operations; (2) measurement; (3) geometry and spatial sense; (4) algebraic thinking; and (5) data analysis and probability. Prior to 2001, the FCAT math score was a composite score that combined student performance on the multiple choice and performance portions of the math test. The 2001 FCAT scale score is based only on the multiple choice portion of the test. The assessment administered in the 2001–02 school year had a reliability ranging from .89 to .93. The author reports the correlation between the 2000 NCE score and 2001 Scale score as 0.512 (as cited in Kerstyn, 2001, 2002, and 2004).

Appendix C: Findings included in the rating for the mathematics achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Kerstyn, 2002[a]								
2001 FCAT Math Scale Score	Grade 8 (MJ-3)	64 classrooms/ 1,420 students	298.00 (38.23)	294.40 (38.09)	3.60	0.09	+4	.20
2001 FCAT Math Scale Score	Grade 8 (MJ-3 Advanced)	20 classrooms/ 430 students	nr	nr	na	na	na	.31
2002 FCAT Math Scale Score	Grade 8 (MJ-3)	328 classrooms/ 5,957 students	293.81 (31.96)	289.90 (40.55)	3.91	0.10	+4	.04
2002 FCAT Math Scale Score	Grade 8 (MJ-3 Advanced)	112 classrooms/ 2,079 students	329.93 (23.30)	331.04 (24.08)	–1.11	–0.05	–2	.42
Domain average for mathematics achievement (Kerstyn, 2002)						**0.05**	**+2**	**Not statistically significant**
Domain average for mathematics achievement across all studies						**0.05**	**+2**	**na**

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 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 WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. FCAT = Florida Comprehensive Assessment Test. na = not applicable. nr = not reported

[a] For Kerstyn (2002) the *p*-values presented here were reported in the original study. A correction for multiple comparisons was needed and resulted in a WWC-computed critical *p*-value of .013 for the 2002 FCAT Math Scale Score for students in MJ-3 classrooms; therefore, the WWC does not find the result to be statistically significant. The 2001 analyses were conducted using classroom-level means and ANCOVA, and the 2002 analyses were conducted using student-level scores and hiercarchical linear modeling (HLM). The study did not report data needed to calculate adjusted mean differences or student-level effect sizes for the 2001 FCAT for MJ-3 Advanced classrooms. The analytic sample for this finding has baseline differences between the intervention and comparison groups of 0.17 standard deviations, which is in the range for which the WWC requires an adjustment for the baseline measure. The author conducted an analysis that adjusted for the baseline measure and reported that the finding was positive and not statistically significant. The means for the 2001 FCAT for MJ-3 classrooms do not match those in the previous 2009 report because the previous report used student-level means, whereas this table reports classroom-level means, consistent with the unit of analysis. The intervention group mean for the 2002 outcomes were obtained by adding the HLM coefficient to the comparison group mean. The WWC excludes findings without an effect size from the domain averages. This study is characterized as having an indeterminate effect because the mean effect reported is neither statistically significant nor substantively important (0.25 standard deviations or larger). For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

Appendix D: Description of supplemental findings for the mathematics achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Kerstyn, 2002[a]								
2001 Cumulative Test	Grade 8 (MJ-3)	66 classrooms/ 1,593 students	31.40 (nr)	30.90 (nr)	0.50	na	na	.61
2001 Cumulative Test	Grade 8 (MJ-3 Advanced)	20 classrooms/ 462 students	nr	nr	na	na	na	.53
2002 FCAT Math Scale Score	Grade 8 (MJ-3 students in 36 middle schools)	258 classrooms/ 4,045 students	294.16 (33.46)	295.06 (33.91)	–0.90	–0.03	–1	.47
2002 FCAT Math Scale Score	Grade 8 (MJ-3 students in 14 middle schools)	113 classrooms/ 1,676 students	294.32 (31.46)	290.48 (35.62)	3.84	0.12	+5	.02

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. 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 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. Some statistics may not sum as expected due to rounding. nr = not reported. na = not applicable.

[a] For Kerstyn (2002), the *p*-values presented here were reported in the original study. A correction for multiple comparisons was needed and resulted in a WWC-computed critical *p*-value of .013 for the 2002 FCAT Math Scale Score for MJ-3 Students in 14 middle schools; therefore, the WWC does not find the result for this outcome to be statistically significant. The intervention group mean for the 2002 outcomes were obtained by adding the HLM coefficient to the comparison group mean. The 2001 outcomes are based on classroom-level means, consistent with the unit of analysis. The study did not report data needed to calculate student-level effect sizes for the 2001 outcomes.The WWC cannot report an adjusted mean difference for the 2001 Cumulative Test for MJ-3 Advanced classrooms, because the data were not reported by the author. The analytic sample has baseline differences between the intervention and comparison groups of 0.16 standard deviations, which is in the range for which the WWC requires an adjustment for the baseline measure. The author conducted an analysis that adjusted for the baseline measure and reported that the finding was positive and not statistically significant. The contrasts in this table include supplementary time periods and analytic samples reported in the study. For more information on these contrasts, see Appendix A.

Endnotes

[*] Due to the 2015 restructuring of the Mathematics topic area from three areas (Elementary, Middle, and High School) to two areas (Primary and Secondary Mathematics), this is considered a new report, rather than an updated report. The information in this report combines the research examined in the prior reports and presents the conclusions differently.

[1] The descriptive information for this intervention comes from a publicly available source—the developer's website (www.icanlearn.com, downloaded November 2016). The What Works Clearinghouse (WWC) provided the developer with the intervention description in December 2016, and asked the developer to review it for accuracy from their perspective. The WWC subsequently incorporated feedback from the developer. Further verification of the accuracy of the descriptive information for this intervention is beyond the scope of this review.

[2] The WWC previously released reports on *I CAN Learn*® under the Middle School Mathematics (MSM) topic area in March 2009 and the High School Mathematics (HSM) topic area in February 2012; the WWC prepared the reports using the WWC Procedures and Standards Handbook (versions 1.0 and 2.0, respectively), and the Middle and High School Mathematics review protocols (versions 1.0 and 2.1, respectively). In June 2015, the WWC restructured the reviews of research on math interventions into two areas instead of three. These two review areas are Primary Mathematics (which includes interventions in which math is presented through multi-topic materials and curricula, typically used in grades K–8), and Secondary Mathematics (which includes interventions organized by math content area [e.g., algebra, geometry, and calculus], typically taught in grades 9–12). These two areas are replacing the prior Elementary School Mathematics, MSM, and HSM topic areas, which were organized by student grade level. The WWC is updating and replacing intervention reports written under the prior topic areas.

The literature search reflects documents publicly available by November 2016. This updated report includes reviews of 14 studies that the previous WWC intervention reports did not include. Of the additional studies, 12 were not within the scope of the review protocol for the Primary Mathematics topic area, and two were within the scope of the review protocol for the Primary Mathematics topic area but did not meet WWC group design standards. A complete list and disposition of all studies reviewed are available in the references.

The current report, which includes reviews of all previous studies that met WWC group design standards with or without reservations, resulted in a revised disposition for 14 studies.

Barrow et al. (2009) is rated *does not meet WWC group design standards* in this report, whereas it had previously received a rating of *meets WWC group design standards with reservations* in the HSM intervention report and a rating of *does not meet WWC group design standards* in the MSM intervention report. The prior reviews focused on results by middle and high school (grade level), and were conducted based on versions 1.0 and 2.1 of the WWC Procedures and Standards Handbook for the MSM and HSM reports, respectively. In both prior reviews, the WWC found the study to have high attrition; therefore, the study needed to demonstrate equivalence. In response to a query for the HSM review, the authors provided evidence of equivalence for the analytic sample (that is, the sample used for study analysis) reviewed for the HSM report. The authors did not respond to a similar WWC query for the MSM report. Therefore, the two prior ratings differed. In the current review based on the revised Primary Mathematics topic area, the review focused on students in pre-algebra courses, regardless of grade level. The WWC sent a question to the authors for attrition information and evidence of baseline equivalence for the eligible sample of students in pre-algebra courses, but the authors did not provide this information; therefore, attrition could not be assessed and baseline equivalence was not demonstrated. As a result, the study does not meet standards.

Kirby (2004a) is rated *does not meet WWC group design standards* in this report, whereas it had previously received a rating of *meets WWC group design standards with reservations* in the MSM intervention report, which was conducted based on version 1.0 of the WWC Procedures and Standards Handbook. There were two reasons for the change in the study rating. First, the prior review assumed that there was no attrition during the school year. The current review did not make that assumption and did not have the necessary information to calculate attrition. The WWC sent a question to the author to request attrition data but did not receive a response. In the absence of attrition information, the WWC requires the study to demonstrate baseline equivalence. However, the study does not present baseline data; therefore, equivalence is not demonstrated as required. Second, the study has a confounding factor because only one teacher was assigned to the intervention condition. Under version 1.0 of the WWC Standards, a study could be rated *meets WWC evidence standards* even with the presence of a confounding factor. Under version 3.0 of the WWC Standards, the study cannot meet standards when a confounding factor is present.

Parrott (2005) is *ineligible for review* in the current report, whereas it was rated *does not meet WWC group design standards* in the previous MSM intervention report based on version 1.0 of the WWC Procedures and Standards Handbook. The rating has changed due to a change in WWC eligibility rules under version 3.0 of the Standards. Under the version 3.0 WWC Standards, master's theses are not eligible for review; therefore, the study is not eligible for the current review.

The remaining 11 studies with rating changes are *ineligible for review* in this report, whereas they were rated differently in a previous report. In all 11 cases, the change in rating is due to the restructuring of the Mathematics topic area from three areas (Elementary, Middle, and High) to two areas (Primary and Secondary). The prior topic areas were defined by grade level, whereas the current topic areas are defined by course content.

(1) Kirby (n.d.) is *ineligible for review* in this report, whereas it was previously rated *does not meet WWC group design standards* in the MSM and HSM intervention reports. The prior reviews were based on versions 1.0 and 2.1 of the WWC Procedures and Standards Handbook, respectively. In the current report, the study is ineligible for review because the intervention examines the effectiveness of *I CAN Learn® Algebra*, which the WWC considers to be a secondary mathematics course and is not eligible for review under the Primary Mathematics topic area.

(2) Kirby (2003) is *ineligible for review* in this report, whereas it was rated *does not meet WWC group design standards* in the previous MSM report. The prior review was conducted based on version 1.0 of the WWC Procedures and Standards Handbook. The study includes eighth grade students and the author does not indicate whether they used *I CAN Learn® Algebra* or *I CAN Learn® Pre-Algebra*. The WWC sent a question to the author to clarify the intervention used but did not receive a response. Because it is unknown whether the study uses an eligible intervention, the study is ineligible for review.

(3) Kirby (2004d) is *ineligible for review* in this report, whereas it was rated *Meets WWC group design standards with reservations* in the previous MSM report. The prior review was based on version 1.0 of the WWC Procedures and Standards Handbook. The study includes eighth grade students and the author does not indicate whether they used *I CAN Learn® Algebra* or *I CAN Learn® Pre-Algebra*. The WWC sent a question to the author to clarify the intervention used but did not receive a response. Because it is unknown whether the study uses an eligible intervention, the study is ineligible for review.

(4) Kirby (2004e) is *ineligible for review* in this report, whereas it was previously rated *does not meet WWC group design standards* in the MSM and *ineligible for review* in the HSM intervention reports. The prior reviews were conducted based on versions 1.0 and 2.1 of the WWC Procedures and Standards Handbook, respectively. In the current report, the study is ineligible for review because the intervention examines the effectiveness of *I CAN Learn® Algebra*, which the WWC considers to be a secondary mathematics course and is not eligible for review under the Primary Mathematics topic area.

(5) Kirby (2005) is *ineligible for review* in this report, whereas it was previously rated *does not meet WWC group design standards* in the MSM and HSM intervention reports. The prior reviews were conducted based on versions 1.0 and 2.1 of the WWC Procedures and Standards Handbook, respectively. In the current report, the study is ineligible for review because the intervention examines the effectiveness of *I CAN Learn® Algebra*, which the WWC considers to be a secondary mathematics course and is not eligible for review under the Primary Mathematics topic area.

(6) Kirby (2006a) is *ineligible for review* in this report, whereas it was previously rated *does not meet WWC group design standards* in the MSM and *ineligible for review* in the HSM intervention reports. The prior reviews were based on versions 1.0 and 2.1 of the WWC Procedures and Standards Handbook, respectively. In the current report, the study is ineligible for review because the intervention examines the effectiveness of *I CAN Learn® Algebra*, which the WWC considers to be a secondary mathematics course and is not eligible for review under the Primary Mathematics topic area.

(7) Kirby (2006b) is *ineligible for review* in this report, whereas it was rated *meets WWC group design standards with reservations* in the previous MSM report. The prior review was based on version 1.0 of the WWC Procedures and Standards Handbook. The study includes eighth grade students and the author does not indicate whether they used *I CAN Learn® Algebra* or *I CAN Learn® Pre-algebra*. The WWC sent a question to the author to clarify the intervention used but did not receive a response. Because it is unknown whether the study uses an eligible intervention, the study is ineligible for review.

(8) Kirby (2006d) is *ineligible for review* in this report, whereas it was previously rated *does not meet WWC group design standards* in the MSM and HSM intervention reports. The prior reviews were based on versions 1.0 and 2.1 of the WWC Procedures and Standards Handbook, respectively. In the current report, the study is ineligible for review because the intervention examines the effectiveness of *I CAN Learn® Algebra*, which the WWC considers to be a secondary mathematics course and is not eligible for review under the Primary Mathematics topic area.

(9) Oescher (2002) is *ineligible for review* in this report, whereas it was rated *does not meet WWC group design standards* in the previous MSM and HSM reports. The prior reviews were based on versions 1.0 and 2.1 of the WWC Procedures and Standards Handbook, respectively. The study includes eighth grade students and the author does not indicate whether they used *I CAN Learn® Algebra* or *I CAN Learn® Pre-Algebra*. The WWC sent a question to the author to clarify the intervention used but did not receive a response. Because it is unknown whether the study uses an eligible intervention, the study is ineligible for review. In addition, Oescher (2002) is considered a related citation for JRL Enterprises, Inc. (n.d.b) in this report.

(10) Oescher and Kirby (2004) is ineligible for review in this report, whereas it was previously rated does not meet WWC group design standards in the MSM and HSM intervention reports. The prior reviews were based on versions 1.0 and 2.1 of the WWC Procedures and Standards Handbook, respectively. In the current report, the study is ineligible for review because the intervention examines the effectiveness of I CAN Learn® Algebra, which the WWC considers to be a secondary mathematics course and is not eligible for review under the Primary Mathematics topic area.

(11) Scafide (2004) is ineligible for review in this report, whereas the study was rated does not meet WWC group design standards in the previous MSM intervention report. The prior review was based on version 1.0 of the WWC Procedures and Standards Handbook. The study includes eighth grade students and the author does not indicate whether they used I CAN Learn® Algebra or I CAN Learn® Pre-Algebra. The WWC sent a question to the author to clarify the intervention used but did not receive a response. Because it is unknown whether the study uses an eligible intervention, the study is ineligible for review.

In addition to the rating changes, this review effort identified some citations to be related and therefore to comprise a single study, whereas the prior reports treated the citations as separate studies. These changes are due to clarification of a study definition in version 3.0 of the WWC Standards, and are notable in terms of the effectiveness rating and WWC-reported findings for Kerstyn (2001), Kerstyn (2002), and Kerstyn (2004), which are considered related citations for one study in this report. In the prior MSM report, the WWC treated Kerstyn (2001) and Kerstyn (2002) as separate studies, and Kerstyn (2004) was not included. These studies are now considered multiple citations from one study, and they collectively received the rating of meets WWC group design standards with reservations, which is the same rating in the current and prior report.

Reviews of studies in this report used the standards from the WWC Procedures and Standards Handbook (version 3.0) and the Primary Mathematics topic area review protocol (version 3.1). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[3] Please see the Primary Mathematics review protocol (version 3.1) for more information about the outcome domain.

[4] For criteria used to determine the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 17. These improvement index numbers show the average and range of individual-level improvement indices for the finding in the study.

[5] Kerstyn (2001, 2002, and 2004) also presented results for Algebra I classrooms that were ineligible for review under the Primary Mathematics topic area; therefore, the Algebra I classrooms are not discussed in this report. The results based on Algebra I classrooms are eligible for review under the Secondary Mathematics topic area.

[6] The study also used an additional outcome, the Cumulative Test, which is a supplemental outcome that does not contribute to the effectiveness rating. Appendix B describes the Cumulative Test.

[7] The second year analyses presented in Appendix D excluded students whose scores do not count toward school accountability grades, such as those categorized as ESE, LEP, or home-schooled, which the author called standard curriculum students. The main findings in Appendix C are from analyses that included all students, which the author called all curriculum students.

Recommended Citation

What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2017, July). *Primary Mathematics intervention report: I CAN Learn®.* Retrieved from https://whatworks.ed.gov

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC group design standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC group design standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC group design standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC group design standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all subjects initially assigned to the intervention and comparison groups. If a randomized controlled trial (RCT) or regression discontinuity design (RDD) study has high levels of attrition, the validity of the study results can be called into question. An RCT with high attrition cannot receive the highest rating of *Meets WWC Group Design Standards without Reservations*, but can receive a rating of *Meets WWC Group Design Standards with Reservations* if it establishes baseline equivalence of the analytic sample. Similarly, the highest rating an RDD with high attrition can receive is *Meets WWC RDD Standards with Reservations*.

For single-case design research, attrition occurs when an individual fails to complete all required phases or data points in an experiment, or when the case is a group and individuals leave the group. If a single-case design does not meet minimum requirements for phases and data points within phases, the study cannot receive the highest rating of *Meets WWC Pilot Single-Case Design Standards without Reservations*.

Baseline A point in time before the intervention was implemented in group design research and in regression discontinuity design studies. When a study is required to satisfy the baseline equivalence requirement, it must be done with characteristics of the analytic sample at baseline. In a single-case design experiment, the baseline condition is a period during which participants are not receiving the intervention.

Clustering adjustment 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 method by which intervention and comparison groups are assigned (group design and regression discontinuity design) or the method by which an outcome measure is assessed repeatedly within and across different phases that are defined by the presence or absence of an intervention (single-case design). Designs eligible for WWC review are randomized controlled trials, quasi-experimental designs, regression discontinuity designs, and single-case designs.

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 and inclusion in this report 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.

Extent of evidence An indication of how much evidence from group design studies supports the findings in an intervention report. The extent of evidence categorization for intervention reports focuses on the number and sizes of studies of the intervention in order to give an indication of how broadly findings may be applied to different settings. There are two extent of evidence categories: small and medium to large.

- **small:** includes only one study, or one school, or findings based on a total sample size of less than 350 students and 14 classrooms (assuming 25 students in a class)
- **medium to large:** includes more than one study, more than one school, and findings based on a total sample of at least 350 students or 14 classrooms

Glossary of Terms

Gain scores The result of subtracting the pretest from the posttest for each individual in the sample. Some studies analyze gain scores instead of the unadjusted outcome measure as a method of accounting for the baseline measure when estimating the effect of an intervention. The WWC reviews and reports findings from analyses of gain scores, but gain scores do not satisfy the WWC's requirement for a statistical adjustment under the baseline equivalence requirement. This means that a study that must satisfy the baseline equivalence requirement and has baseline differences between 0.05 and 0.25 standard deviations *Does Not Meet WWC Group Design Standards* if the study's only adjustment for the baseline measure was in the construction of the gain score.

Group design A study design in which outcomes for a group receiving an intervention are compared to those for a group not receiving the intervention. Comparison group designs eligible for WWC review are randomized controlled trials and quasi-experimental designs.

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.

Intervention An educational program, product, practice, or policy aimed at improving student outcomes.

Intervention report A summary of the findings of the highest-quality research on a given program, product, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against design standards, and summarizes the findings of those that meet WWC design standards.

Multiple comparison adjustment An adjustment to the statistical significance of results to account for multiple comparisons in a group design study. The WWC uses the Benjamini-Hochberg (BH) correction to adjust the statistical significance of results within an outcome domain when study authors perform multiple hypothesis tests without adjusting the p-value. The BH correction is used in three types of situations: studies that tested multiple outcome measures in the same outcome domain with a single comparison group; studies that tested a given outcome measure with multiple comparison groups; and studies that tested multiple outcome measures in the same outcome domain with multiple comparison groups. Because repeated tests of highly correlated constructs will lead to a greater likelihood of mistakenly concluding that the impact was different from zero, in all three situations, the WWC uses the BH correction to reduce the possibility of making this error. The WWC makes separate adjustments for primary and secondary findings.

Outcome domain A group of closely-related outcomes. A domain is the organizing construct for a set of related outcomes through which studies claim effectiveness.

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.

Rating of effectiveness For group design research, the WWC rates the effectiveness of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. For single-case design research, the WWC rates the effectiveness of an intervention in each domain based on the quality of the research design and the consistency of demonstrated effects. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 17.

Regression discontinuity design (RDD) A design in which groups are created using a continuous scoring rule. For example, students may be assigned to a summer school program if they score below a preset point on a standardized test, or schools may be awarded a grant based on their score on an application. A regression line or curve is estimated for the intervention group and similarly for the comparison group, and an effect occurs if there is a discontinuity in the two regression lines at the cutoff.

Single-case design 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 tend to be spread out over a large range of values.

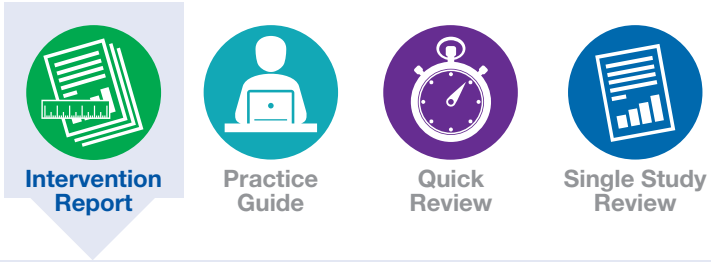
Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p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

Study rating The result of the WWC assessment of a study. The rating is based on the strength of the evidence of the effectiveness of the educational intervention. Studies are given a rating of *Meets WWC Design Standards without Reservations*, *Meets WWC Design Standards with Reservations*, or *Does Not Meet WWC Design Standards*, based on the assessment of the study against the appropriate design standards. The WWC has design standards for group design, single-case design, and regression discontinuity design studies.

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

Systematic review A review of existing literature on a topic that is identified and reviewed using explicit methods. A WWC systematic review has five steps: 1) developing a review protocol; 2) searching the literature; 3) reviewing studies, including screening studies for eligibility, reviewing the methodological quality of each study, and reporting on high quality studies and their findings; 4) combining findings within and across studies; and 5) summarizing the review.

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



An **intervention report** summarizes the findings of high-quality research on a given program, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against evidence standards, and summarizes the findings of those that meet standards.

This intervention report was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.