What Works Clearinghouse

Detailed Study Report

Reviewed Study: **Kerstyn, C. (2001). *Evaluation of the I CAN LEARN Mathematics Classroom. First year of implementation (2000–2001 school year).* Unpublished manuscript.**

WWC Study Reports are intended to support decision making; neither the What Works Clearinghouse (WWC) nor the U.S. Department of Education endorses any interventions. No single Study Report should be used as a basis for making policy decisions because (1) few studies are designed and implemented flawlessly and (2) all studies are tested on a limited number of participants, using a limited number of outcomes, at a limited number of times, so generalizing from one study to any context is very difficult. To highlight these issues, the WWC Study Reports describe in detail the specifics of each study, focusing primarily on studies that provide the best evidence of effects (randomized controlled trials). Systematic reviews of the evidence will be conducted to extend the results of the individual studies.

Brief version of the report (PDF)

Topic:	Curriculum-Based Interventions for Increasing K–12 Math Achievement—Middle School
Intervention:	I Can Learn
Research Design:	Quasi-Experimental Design with Matching
Study Rating:	
Date Released:	June 30, 2004

= Meets Evidence Standards = Meets Evidence Standards with Reservations = Does Not Meet Evidence Standards

The What Works Clearinghouse (www.whatworks.ed.gov) was established in 2002 by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. Please email all questions and comments to info@whatworks.ed.gov. The What Works Clearinghouse is administered by the U.S. Department of Education through a contract to a joint venture of the American Institutes for Research and the Campbell Collaboration.

WWC Study Ratings[a]
Kerstyn (2001)

Causal Validity: Meets WWC Evidence Standards with Reservations, Quasi-Experimental Design with Matching

Participants in the intervention group were matched on several relevant variables to similar participants in the control group. There were no significant differences between groups on a pretest of mathematics achievement. There was minimal attrition (one classroom and its matched control were dropped because of a midyear change in curriculum), and no extraneous events were identified that appeared to be confounded with the intervention's effect.

Other Study Characteristics	Study Rating	Study-Specific Information
Intervention Fidelity	●	The I Can Learn (ICL) curriculum reflects commonly held or theoretical conceptions of the characteristics that such an intervention should contain. The author also provides enough information about the curriculum, the population, and the setting to allow replication. However, teacher surveys indicate that teachers differed in their implementation of the curriculum.
Outcome Measures	●●	Although five achievement outcomes were presented in this study, only two are the target outcome measures in this report and meet the criteria for this standard. The target outcomes are measures of math achievement, which is the content of interest for this report, and are properly aligned.
People, Settings, and Timing	●	Although some important characteristics are represented in the sample, many are not. The sample of students was part of the identified population, but it included variation on only some of the important student characteristics, namely, gender, race, and socioeconomic status.
Testing within Subgroups	●	The intervention's effect was tested across the entire sample but not within important subgroups, except for tracking.
Analysis	●	The unit of assignment (class) was the same as the units of analysis and intervention delivery. In this study, natural student groupings (classrooms, schools, etc.) may have affected findings. Although the author's analysis does not address this grouping problem, the author does not report significant positive findings, so the impact of groupings on findings is likely minimal. The statistical properties of the data allowed for valid estimates of the effect sizes. However, the sample sizes were not adequate at the class level to allow for sufficiently precise estimates of the effect size.
Statistical Reporting	●●	The statistical tests were adequately reported, and effect sizes could be estimated for the outcome measure of interest.

Summary of Results. There were no statistically significant differences between the intervention and control groups on the two target outcome measures analyzed in this report. Different levels of implementation may have affected this finding.

Note. ●● Fully meets criteria; ● Meets minimum criteria; **X** Does not meet criteria.

[a] For more information on the criteria used to rate this study, see the WWC Study Review Standards.

Intervention: I Can Learn

Operational Features

I Can Learn Algebra (ICL) is a math curriculum for use in grades 7 through 10 that was developed by New Orleans-based JRL Enterprises. ICL is a software-based math curriculum that the developer indicates meets the National Council of Teachers of Mathematics standards. The ICL software is designed to be interactive and to allow students to progress through lessons at their own pace. Teachers are expected to play a role in determining the content of the lesson and other aspects of the class. At the beginning of the year, the teacher determines many factors about the structure of the course such as homework assignments, lesson organization, lesson presentation, manipulatives used, assessments, and grade evaluations.

Each ICL lesson follows a five-part format consisting of pretest, review, lesson presentation, quiz, and cumulative review. The pretest covers material from the upcoming lesson. If students miss one question on the pretest, they continue into the current lesson. If students get all the questions right on the pretest, they may advance on to the next lesson if the teacher enables the software's advancement option. The developer describes the lessons as being grouped together like the lessons of a chapter in a textbook. Students take a cumulative review of the concepts taught after they complete a lesson.

In this study, there were two conditions: classrooms in schools that were implementing ICL were matched (that is, on instruction time, class mean prior achievement, size of class, percentage of students on free and reduced-price lunch, percentage of minority enrollment, and time of day) with classrooms that were using a traditional math curriculum.

People, Settings, and Timing

The ICL curriculum was designed for ethnically diverse, inner-city students in grades seven through 10. The target population in this study was 8th-grade students in Title I middle schools within the Hillsborough County Public School system in Florida. This county includes the Tampa metro area. The study was limited to regular education students. Students were racially diverse and many were eligible for free or reduced-price lunch.

Cost Information

The author does not provide the cost of implementing the ICL curriculum in the district.

Intended Duration

The author indicates that ICL consists of 109 lessons. However, Kerstyn does not indicate how many of the 109 lessons are required to be completed in order for the curriculum to be implemented as intended. The author also indicates that the ICL curriculum was implemented in class periods of 45, 50, 80, and 90 minutes in length. The author does not indicate how long a class period the curriculum is intended to cover. When surveyed, the teachers reported that 45 minute classes were not long enough to make it through the curriculum.

Scientific Rationale

The author does not provide any scientific rationale for conducting the study except to say that the school district entered into a three-year contract with the distributor of the ICL curriculum and wanted to evaluate whether using the curriculum had a positive effect on student achievement.

Overview of the Study

Purpose

The central question of this study was whether there was a difference in mathematics achievement between students taught in an ICL classroom and those taught in a traditional classroom. The author was also interested in the experience students had while using the software. Thus, students in the ICL curriculum were surveyed about their attitudes towards using the computer lab for instruction. Teachers were surveyed about their opinions regarding the use of the ICL program and which

instructional practices they found most useful. Parents of students in the ICL classroom were surveyed on their opinions of their child's progress in mathematics class over the year of the study. This WWC Study Report focuses only on the achievement component of the study.

Intervention Fidelity

The author reports some problems with the implementation of the program resulting from the lack of guidelines for teachers regarding implementation. The teachers in this study participated in training sessions on the use of the software and hardware, but not on the use of the software in instruction. Kerstyn believes that because of the lack of guidelines, there were differences in instructional practices between teachers.

Implementation was monitored in three ways;

1. The ICL teachers discussed instructional and technical concerns with each other in meetings throughout the study year.
2. Some ICL classrooms were videotaped; the videotapes were to be used in the training of new ICL teachers.
3. All of the ICL teachers were asked to fill out a survey at the end of the year on which they were asked to comment on how the curriculum was implemented.

Through these monitoring systems, the author determined that teachers differed in their implementation of the ICL program. She indicates that this may have affected the results of the study.

Outcome Measures

There were five math achievement outcomes in this study. Four of the outcomes—the Algebra I semester exam, Algebra I Honors semester exam, the MJ-3 Cumulative Test, and the MJ-3 Advanced Cumulative Test—are end-of-semester exams that are administered after the end of the first semester of the program. The district created these exams, and Kerstyn indicates that they had undergone rigorous evaluation before their administration. However, the author only presents the reliability information for one of these tests, the MJ-3 Cumulative Test (Kerstyn, private correspondence). Therefore, this study report does not discuss findings for the other three end-of-semester exams without reliability information.

The fifth outcome is the Florida Comprehensive Assessment Test (FCAT), which was administered in February 2001. The author does not present the reliability information for this test; however, this information is available in a technical report written by the Florida Department of Education (2002). This WWC Study Report focuses only on the FCAT measures, because this assessment was taken by all students and is the only assessment with independently documented reliability and validity information.

Research Design

The research design for this study was quasi-experimental with matching. Classrooms were nonrandomly assigned to the intervention and control groups. The study author included 8th-grade classes from 36 middle schools in the Hillsborough County School District in the sample. The researcher matched ICL classrooms with control classrooms on the basis of instructional time, prior achievement, class size, proportion of students on free or reduced-price lunch, proportion of minority students, and time of day. Eight pairs of Algebra I classes, eight pairs of Algebra I Honors classes, 10 pairs of MJ-3 advanced (Advanced Prealgebra) classes, and 33 pairs of MJ-3 (Prealgebra) classes resulted from the matching. The MJ-3 course is the standard 8th-grade mathematics course, whereas MJ-3 Advanced, Algebra I, and Algebra I Honors are advanced 8th-grade mathematics courses. Initially there were 118 classes, 59 in each condition. The author indicates only that ICL classrooms were matched with non-ICL classrooms. The author does not provide information on the schools from which the control classrooms were drawn.

The author does not provide any details about which recruitment methods were used, or who was responsible for enrolling and assigning participants.

The unit of assignment (classroom) matches the unit of analysis. Classes were nested within schools, with 3.3 classes on average per school participating in the study.

Participant Flow

The author reports that during the course of the year, one classroom switched from the control to the ICL group. As a result, this classroom and its corresponding match were removed. This was the only case of attrition that the author reported. At the end of the study, 58 pairs of classes remained. (See Figure 1.)

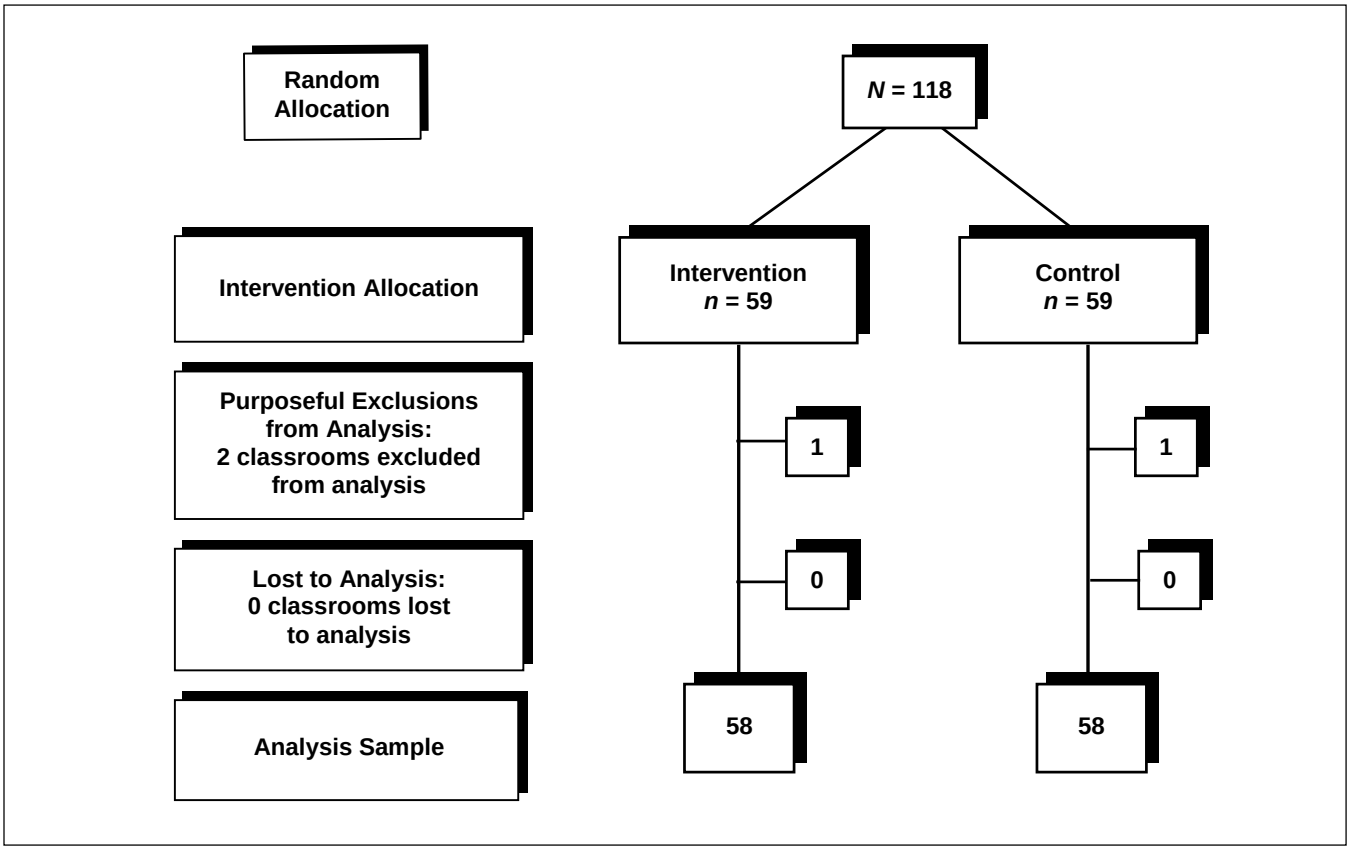
Reference Periods

The author indicates that the schools entered into a contract with JRL Enterprises during the 1999–2000 school year. The author does not indicate whether the sample was selected during that year or the summer before the 2000–2001 school year when the study took place. Baseline data came from the 2000 FCAT Norm Referenced Test Normal Curve Equivalents. Outcomes were measured at the end of the first semester of the 2000–2001 school year and in February 2001.

Baseline Data

The only sample baseline characteristics presented by the author are pretest scores on the FCAT Norm Referenced Test and the fact that all the students were in 8th grade. The author also provides demographic data on the sample that had both a pretest and posttest score**.** The author used the pretest as a basis for matching and indicates that the scores were "comparable." The significance of the difference between the scores is not reported. (See Table 1.)

Figure 1. Participant Flow[a]



[a] Participant flow reflects classrooms, not individual students.

Table 1. Pretest Characteristics of the Study Sample

Class	Intervention group (n = 59)	Control group (n = 59)
Pretest mean score on math achievement: FCAT		
Algebra	69.7	70.5
Algebra I Honors	81.8	82.2
MJ-3	41.6	41.6
MJ-3 advanced	59.4	58.0

Note. FCAT = Florida Comprehensive Assessment Test.

Statistical Methods

One-way analyses of covariance (ANCOVAs) were used to test the significance of the difference between the groups while taking into account pretest scores on the FCAT. Analyses were done separately for each of the four class types: Algebra I, Algebra I Honors, MJ-3, and MJ-3 advanced. The author presents posttest means and standard deviations along with the *F* statistics from the ANCOVA analyses. No evidence of weighting is presented.

Outcomes and Estimation

Tables 2 through 5 and Figures 2a through 2e present the effects of the ICL curriculum on math achievement, as reported by the study author. The tables present adjusted posttest means and standard deviations. In this study, natural student groupings (classrooms, schools, etc.) may have affected findings. Although the author's analysis does not address this grouping problem, the author does not report significant positive findings, so the impact of groupings on findings is likely minimal.

The author reports findings for each class type separately and does not report combined results, so the results are separated by class type below. Kerstyn also reports the results of the surveys of the ICL students, teachers, and parents; however, this report focuses on the achievement results only.

The WWC has not compared effect sizes for this study report because the possible effects of grouping on findings must first be addressed.

Table 2. Impact Reported by Kerstyn (2001): Algebra I

	Intervention group mean score (*SD*) (n = 8)	Control group mean score (*SD*) (n = 8)	Intervention group estimated effect size (and significance)
Posttest on math achievement: FCAT	351.1 (15.6)	345.4 (11.4)	NR[a]

Note. FCAT = Florida Comprehensive Assessment Test.

[a] The author does not report effect sizes, but reports that the difference was not significant.

Table 3. Impact Reported by Kerstyn (2001): Algebra I Honors

	Intervention group mean score (*SD*) (n = 8)	Control group mean score (*SD*) (n = 8)	Intervention group estimated effect size (and significance)
Posttest on math achievement: FCAT	374.2 (11.1)	373.1 (20.5)	NR[a]

Note. FCAT = Florida Comprehensive Assessment Test.

[a] The author does not report effect sizes, but reports that the difference was not significant.

Table 4. Impact Reported by Kerstyn (2001): MJ-3

	Intervention group mean score (*SD*)	Control group mean score (*SD*)	Intervention group estimated effect size (and significance)
Posttest on math achievement: FCAT	298.0 (15.6)	294.4 (13.8)	NR[a]
Posttest on math achievement: MJ-3 Cumulative Test	31.4 (4.6)	30.9 (5.1)	NR[a]

Note. FCAT = Florida Comprehensive Assessment Test. Preattrition/postattrition sample size (classes) was 33/32 for both the intervention and control groups. There was no attrition for the MJ-3 exam (33/33).

[a] The author does not report effect sizes, but reports that the difference was not significant.

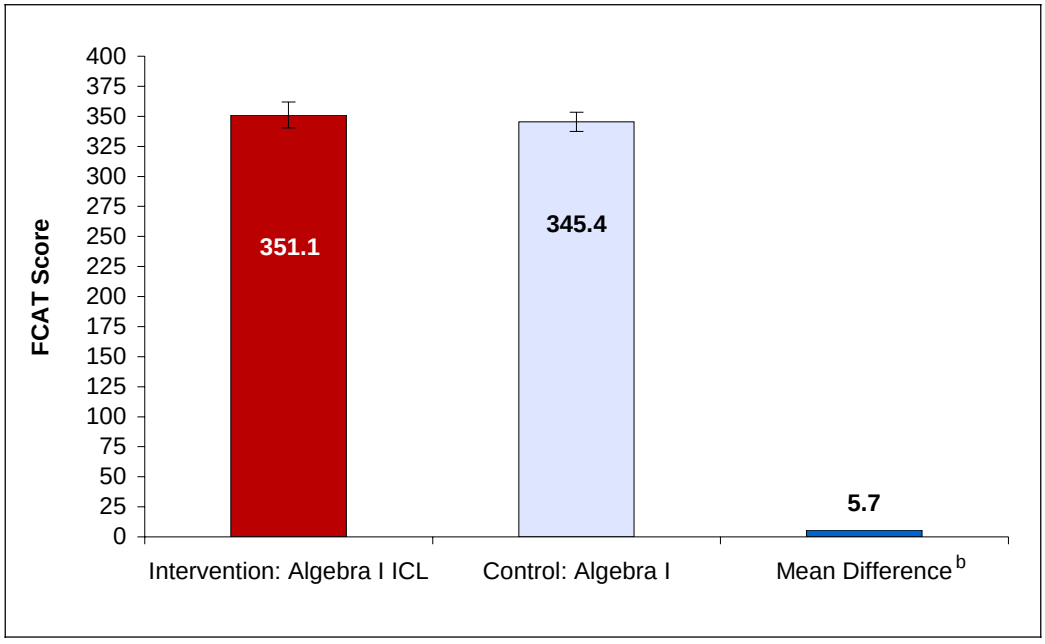
Table 5. Impact Reported by Kerstyn (2001): MJ-3 Advanced

	Intervention group mean score (*SD*) (*n* = 10)	Control group mean score (*SD*) (*n* = 10)	Intervention group estimated effect size (and significance)
Posttest on math achievement: FCAT	331.5 (12.6)	326.1 (11.0)	NR[a]

Note. FCAT = Florida Comprehensive Assessment Test.

[a] The author does not report effect sizes, but reports that the difference was not significant.

Figure 2a. Impact Reported by Kerstyn (2001):[a] Algebra I

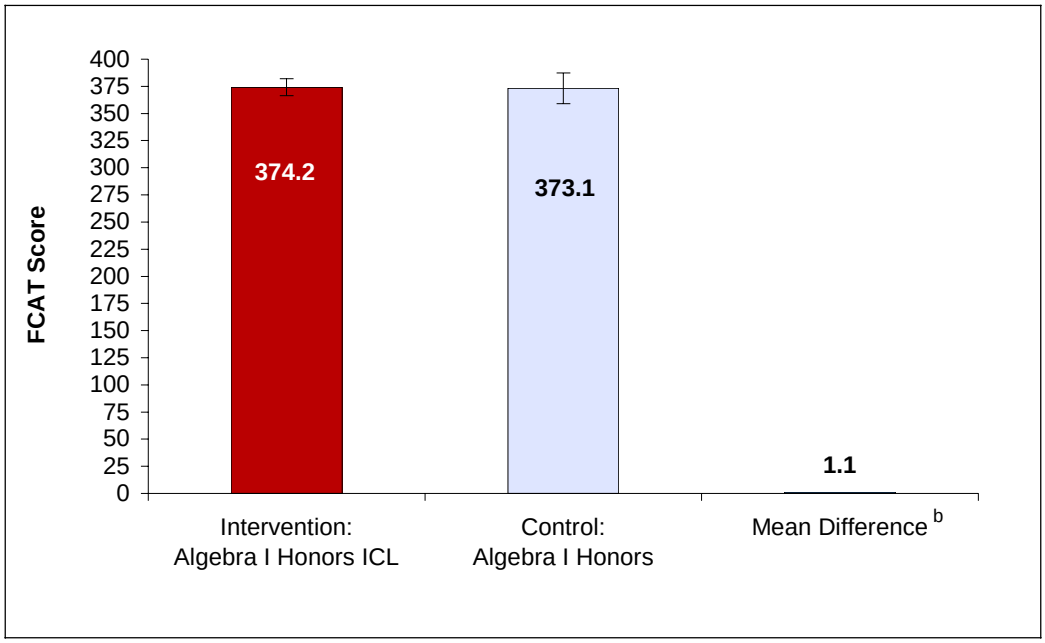


Note. FCAT = Florida Comprehensive Assessment Test.

[a] Confidence intervals were computed by the WWC.

[b] The intervention group scores were not significantly different from the control group scores.

Figure 2b. Impact Reported by Kerstyn (2001):[a] Algebra I Honors

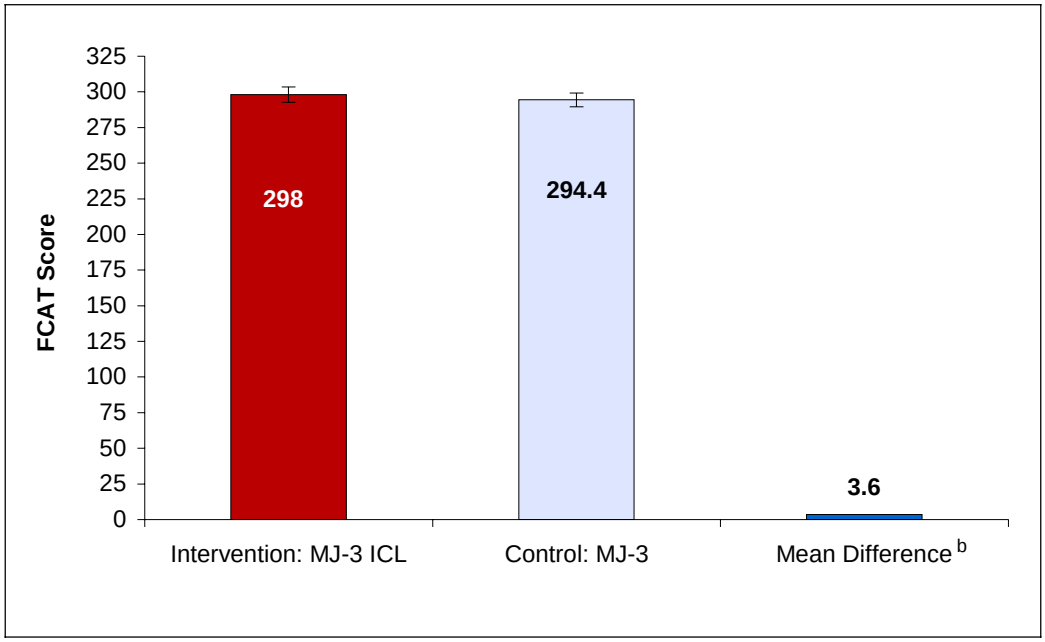


Note. FCAT = Florida Comprehensive Assessment Test.

[a] Confidence intervals were computed by the WWC.

[b] The intervention group scores were not significantly different from the control group scores.

Figure 2c. Impact Reported by Kerstyn (2001):[a] MJ-3

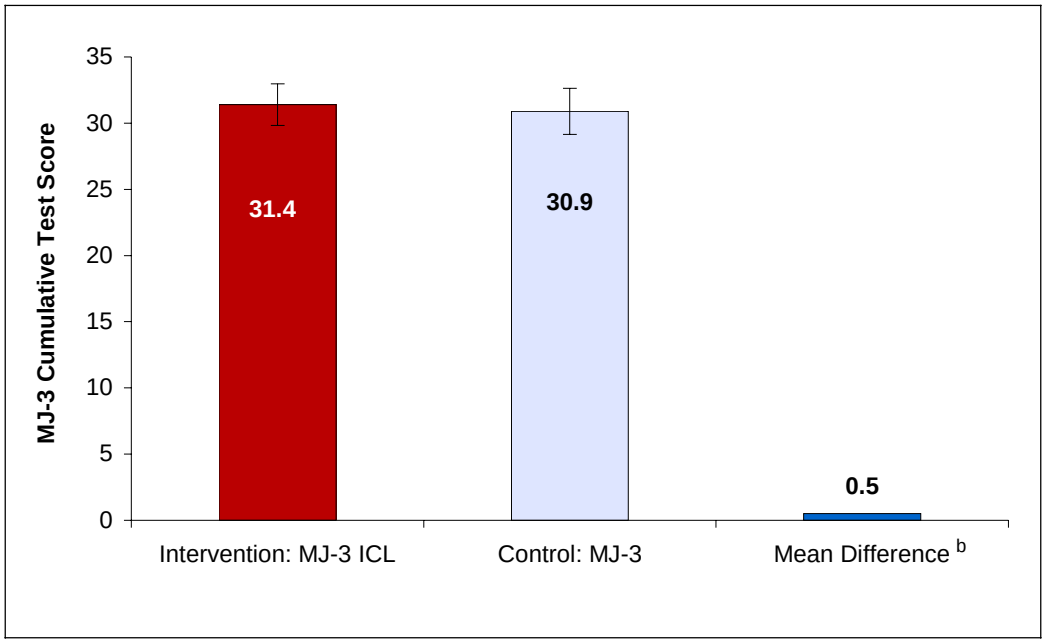


Note. FCAT = Florida Comprehensive Assessment Test.

[a] Confidence intervals were computed by the WWC.

[b] The intervention group scores were not significantly different from the control group scores.

Figure 2d. Impact Reported by Kerstyn (2001):[a] MJ-3

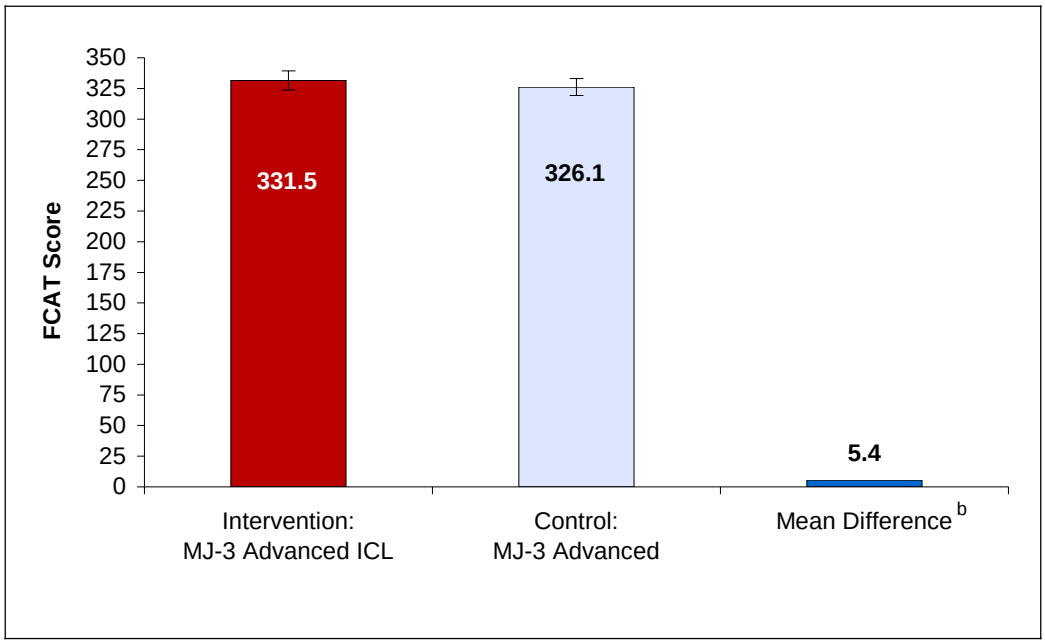


Note. MJ-3 Cumulative Test.

[a] Confidence intervals were computed by the WWC.

[b] The intervention group scores were not significantly different from the control group scores.

Figure 2e. Impact Reported by Kerstyn (2001):[a] MJ-3 Advanced



Note. FCAT = Florida Comprehensive Assessment Test.

[a] Confidence intervals were computed by the WWC.

[b] The intervention group scores were not significantly different from the control group scores.

WWC Evidence Criteria: Middle School Math[a]

Causal Validity

The WWC evidence criteria for determining the level of evidence of a study reviewed under the topic Middle School Math Interventions are:

Meets Evidence Standards

- Randomized controlled trial with no randomization,[b] attrition, or disruption problems
- Regression discontinuity study with no comparability, attrition, or disruption problems

Meets Evidence Standards with Reservations

- Randomized controlled trial with a randomization,[c] attrition, and/or disruption problem
- Regression discontinuity study with a comparability, attrition, or disruption problem
- Quasi-experimental design with equivalent groups and no problems with attrition or disruption

Other Study Characteristics

In addition to determining whether a study Meets Evidence Standards or Meets Evidence Standards with Reservations, the WWC also assesses the strength of a study's evidence based on the following other study characteristics:

Intervention Fidelity. A study fully meets criteria for Intervention Fidelity (●●) if the intervention contains most of the key characteristics that commonly define it, the author provides evidence of good implementation, and the intervention is documented well enough for others to replicate it. A study meets the minimum criteria (●) if the author does not evaluate implementation or finds partial implementation, or the intervention is not documented. A study is excluded from the review (**X**) if it does not meet the initial screening requirements for the intervention by omitting key characteristics of Middle School Math.

Outcome Measures. A study fully meets criteria for Outcome Measures (●●) if the outcome measure has face validity and reliability, and is not too closely aligned[d] to the content of the intervention. A study meets the minimum criteria (●) if the outcome measure is not too closely aligned to the content of the intervention. A study is excluded from the review (**X**) if it does not meet initial screening requirements by not focusing on important Middle School Math outcomes or if it lacks face validity and/or reliability.

People, Settings, and Timing. A study fully meets criteria for People, Settings, and Timing (●●) if it broadly samples from the people (units of interest) and settings that are the target of the intervention and the outcomes are measured at an appropriate time. A study meets the minimum criteria (●) if narrow but relevant samples and settings are included. A study is excluded from the review (**X**) if it does not include at least a relevant narrow sample of people or settings.

Testing within Subgroups. A study fully meets criteria for Testing within Subgroups (●●) if it identifies important subgroups among its sample and settings, and tests the intervention effect within each subgroup separately. A study meets the minimum criteria (●) if it simply tests the intervention effect across the entire sample. A study is not excluded from the review based on this standard.

Analysis. A study fully meets criteria for Analysis (●●) if the analysis is conducted at the same level (for example, students, classes, schools) as the unit of assignment and the unit of intervention delivery or if there is a mismatch between units but sufficient information is provided to permit an approximate estimation of the intervention's effect and in either case, the data characteristics support the analysis. The study meets the minimum criteria (●) if an approximate estimation of effect at the level of assignment cannot be made. A study is not excluded from the review based on this standard.

Statistical Reporting. A study fully meets criteria for Statistical Reporting (●●) if the findings are reported for most outcome measures and effect sizes can be calculated. The study meets the minimum criteria (●) if findings are reported and effect sizes can be calculated for only some outcome measures. A study is excluded from the review (**X**) if it does not report findings for any relevant outcome measures.

Note. For each study characteristic, the WWC considers a number of features to determine if the study fully meets criteria of that characteristic (●●), meets minimum criteria (●), or does not meet minimum criteria (**X**).

[a] These criteria are applied to studies that have passed initial WWC screening for Middle School Math. For more information on screening requirements.

[b] Studies with randomization problems that make statistical adjustments Meet Evidence Standards.

[c] Studies with randomization problems that do not make statistical adjustments Meet Evidence Standards with Reservations.

[d] An overaligned outcome measure uses material that was part of the intervention. The control group was not exposed to this material.

Reference

Florida Department of Education. (2002). *Technical report: For operational test administrations of the 2000 Florida Comprehensive Assessment Test*. Retrieved on April 7, 2004, from http://www.firn.edu/doe/sas/fcat/pdf/fc00tech.pdf.

Related Studies

To see reports on other studies of I Can Learn.

How Can You Find Out More?

To learn more about this study, read the original study (PDF).

Report Production

Date created: June 30, 2004

Topic area reviewed under: Curriculum-Based Interventions for Increasing K–12 Math Achievement—Middle School.