What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

April 2014

WWC Review of the Report "The Effects of Cognitive Strategy Instruction on Math Problem Solving of Middle School Students of Varying Ability"[1]

The findings from this review do not reflect the full body of research evidence on *Solve It!*

What is this study about?

The study examined the effects of *Solve It!*, a program intended to improve the problem-solving skills of seventh-grade math students. These skills are intended to improve reading and math achievement.

For this 1-year study, researchers matched 40 K–8 and middle schools from Miami–Dade County Public Schools in Florida on Florida Comprehensive Assessment Test (FCAT) scores and socioeconomic status. A school from each pair was randomly assigned to either an intervention group or a comparison group. Administrators in participating schools nominated one seventh-grade math teacher (and his or her students) to participate in the study.

Intervention teachers incorporated the *Solve It!* program into the regular district mathematics curriculum in their math or pre-algebra class periods. Teachers in the comparison group delivered only the regular district mathematics curriculum.

The analysis sample consisted of 1,059 students who returned consent forms (out of approximately 2,500 students initially assigned to participate) and included 644 students from 16 schools in the intervention group and 415 students from 18 schools in the comparison group.

Study authors examined the impact of *Solve It!* on math and reading achievement using math and reading scores from the FCAT. The other outcome assessed in this study—a *Solve It!* curriculum-based assessment—did not meet WWC evidence standards (see Appendix B table notes).[2]

Features of *Solve It!*

Solve It! is an instructional program that explicitly teaches students cognitive strategies of varying complexity to solve math word problems. The goal is for students to internalize the strategies and use them automatically during problem solving. According to the study authors, the following cognitive processes and activities are the focus of instruction: reading and paraphrasing the problem, visualizing the relationship among components, hypothesizing about solutions, estimating and computing the answer, and checking the process and solution.

WWC Rating

The research described in this report meets WWC group design standards with reservations

This study is a randomized controlled trial with high levels of student attrition, but the study demonstrates the equivalence of the analytic sample at baseline on FCAT math and reading scores, and therefore, is able to meet WWC group design standards with reservations. However, the evidence from the *Solve It!* curriculum-based assessment of mathematics does not meet standards because this assessment was overaligned with the content delivered to the intervention group.

What did the study find?

The study authors found no impact of the *Solve It!* intervention on students' math or reading achievement, and the WWC review of the study confirms the non-significant findings for both outcomes.

Appendix A: Study details

Montague, M., Krawec, J., Enders, C., & Dietz, S. (in press). The effects of cognitive strategy instruction on math problem solving of middle school students of varying ability. *Journal of Educational Psychology*.

Setting

The study was conducted in 40 K–8 and middle schools in Miami–Dade County Public Schools.

Study sample

Out of the 78 K–8 and middle schools in Miami–Dade County Public Schools, principals from 40 of the schools agreed to participate in the study. Researchers formed 20 matched pairs of schools based on FCAT scores and socio-economic status, and one school from each pair was randomly assigned to the intervention group. Within each study school, an administrator selected one seventh-grade math teacher (and his or her students) to participate in the intervention.[3] Participating teachers in both conditions had to (a) be "high-quality," (b) be certified to teach math, and (c) teach at least two class periods that enrolled students who were low-achieving and identified as having a learning disability.

Of the 40 schools that initially agreed to participate in the study, six left the study sample, resulting in a final sample of 16 intervention and 18 comparison schools. The analytic sample included 1,059 students from 34 schools who returned consent forms (from a total of about 2,500 students in study classrooms). Overall, 43% of students in the sample were male; 65% were Hispanic, 30% were African American, and 5% were White. Seventy-nine percent of students in the sample were eligible for free or reduced-price lunch.[4]

Intervention group

Students in the intervention group received instruction in *Solve It!* during their regular math class periods beginning in October and continuing for the remainder of the school year. The *Solve It!* instruction was embedded in the regular district curriculum. The intervention began with three scripted lessons and continued with weekly problem solving sessions for the duration of the intervention. On average, teachers delivered 11 *Solve It!* practice sessions during their regular math class periods. Students in the intervention group were assessed seven times during the intervention using a *Solve It!* curriculum-based measure. For the third and subsequent assessments, students and teachers received graphs illustrating students' growth.

Teachers in the intervention group differed on some key characteristics from teachers in the comparison group. Notably, more teachers in the intervention group had doctoral degrees (12% vs. 5%) and fewer were in their first 3 years of teaching (47% vs. 26%).

Comparison group

Students in the comparison group received their regular instruction in the district math curriculum. Teachers in the comparison condition were asked to deliver one lesson on math problem solving each week.

Outcomes and measurement

FCAT math and reading scores were used to assess students' math and reading achievement levels. For a more detailed description of these outcome measures, as well as the other outcome that did not meet standards, see Appendix B.

Support for implementation

The intervention teachers were required to attend a 3-day professional development workshop before school began. Intervention materials were provided to teachers at the beginning of the school year, and practice problems were provided bi-weekly throughout the school year.

Reason for review

This study was identified for review by the WWC because it was supported by a grant to the University of Miami (Principal Investigator: Marjorie Montague) from the Institute of Education Sciences (IES).

Appendix B: Outcome measures for each domain

Domain name	
FCAT Math	The FCAT Grade 7 Math exam assesses ratios and proportional relationships, geometry and measurement, numbers and operations, and base ten numbers.
FCAT Reading	The FCAT Grade 7 Reading exam assesses vocabulary, reading, literary analysis, and the identification of specific information in reading passages.

Table Notes: FCAT Math and Reading results were not included in the final version of the published paper, but were included in a draft version previously reviewed by the WWC. The report also included results for a curriculum-based measure, which is omitted from this WWC report because it is based on test items from the *Solve It!* Manual and is over-aligned with the intervention condition. Additionally, the curriculum-based measure was administered more often to the intervention group (seven times) than to the comparison group (four times).

Appendix C: Study findings for each domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Math achievement								
FCAT Math	7th grade	1,059 students	nr	nr	nr	nr	nr	0.10
Domain average for math achievement						**nr**	**nr**	**Not statistically significant**
Reading achievement								
FCAT Reading	7th grade	1,059 students	299.61 (44.30)	296.77 (41.61)	2.84	0.07	+3	nr
Domain average for reading achievement						**0.08**	**+3**	**Not statistically significant**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students 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 student's percentile rank that can be expected if the student is given the intervention. nr = not reported.

Study Notes: The intervention and comparison group means, standard deviations, and sample sizes used in the WWC review of this article and reported in this table were provided in an author query.

The analysis of the *FCAT Math* outcome required adjustment due to the high levels of sample attrition observed for this contrast and the magnitude of the difference between the intervention and comparison group at baseline. In a prior version of the article, the author reported a non-significant impact on the rate of change between the intervention and comparison groups on this outcome (and this analysis adjusted for baseline differences and for the clustered design of this study). As such, the *p*-value reported for the *FCAT Math* in this WWC report is based on the earlier analysis that includes an adjustment for baseline differences. However, this growth model analysis reported in the earlier version of the paper does not provide sufficient information for the WWC to compute differences in posttest scores, WWC effect sizes, or improvement indices. The information obtained from the authors about unadjusted baseline, posttest means, and standard deviations does not allow for an estimate of program effectiveness that meets WWC standards (the WWC requires an adjustment for baseline differences in a study with high levels of sample attrition and baseline differences in the adjustment range). Therefore posttest means, standard deviations, and WWC calculations of mean difference, effect size, and improvement index for this outcome are not reported in this report.

Because there were small baseline differences (that did not require a statistical adjustment) for the *FCAT Reading* outcome, the WWC calculated the intervention group mean using a difference-in-differences approach (see the WWC Handbook) by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Procedures and Standards Handbook, version 3.0 for more information. The authors' analysis of the *FCAT Reading* outcome did not include point-in-time estimates of the impact of the intervention, so the WWC computed the *p*-value of the impact following the intervention, including a correction for clustering. The authors reported a non-significant *p*-value from an analysis of the difference in rate of student growth between conditions, which does not align with the point-in-time estimate reported by the WWC. The WWC-calculated *p*-value for the point-in-time estimate is also non-significant ($p = 0.68$).

This study is characterized as having no discernible effect on math or reading achievement because no effects are statistically significant or substantively important. For more information, please refer to the WWC Standards and Procedures Handbook, version 3.0, page 26.

Endnotes

[1] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the authors) to assess whether the study design meets WWC group design standards. The review reports the WWC's assessment of whether the study meets WWC group design standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the single study review protocol, version 2.0. The WWC rating applies only to the results that were eligible under this topic area and met WWC group design standards without reservations or met WWC group design standards with reservations, not necessarily to all results presented in the study.

[2] There was one outcome included in the study that is not described in this WWC report. See the table notes in Appendix B for more information.

[3] Because the selection of teachers who participated in the study occurred after random assignment, there is a possibility that the intervention may have influenced teacher decision to participate, and thus, this selection process may have compromised the random assignment procedure. In this study, however, there are already high levels of sample attrition, and therefore, the issue of the possibility of teacher selection is moot for this WWC review. We mention this issue in the design of the study for transparency in the WWC review process.

[4] The demographic characteristics reported in Appendix A were calculated by the WWC by pooling the prevalence rates reported by condition in the study.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014, April). *WWC review of the report: The effects of cognitive strategy instruction on math problem solving of middle school students of varying ability.* Retrieved from http://whatworks.ed.gov

Glossary of Terms

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

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

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

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

Domain A domain is a group of closely related outcomes.

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

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

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

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

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

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects 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 investigators randomly assign eligible participants into intervention and comparison groups.

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

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

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

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

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