WWC Single Study Review
A review of the design and summary of findings for an individual study

ies INSTITUTE OF EDUCATION SCIENCES

June 2017

WWC Review of the Report "Live Webcam Coaching to Help Early Elementary Classroom Teachers Provide Effective Literacy Instruction for Struggling Readers: The Targeted Reading Intervention"[1,2]

The findings from this review do not reflect the full body of research evidence on the *Targeted Reading Intervention (TRI)*.

What is this study about?

The study authors examined the effects of the *Targeted Reading Intervention (TRI)*—a professional development intervention designed to help teachers develop the skills of struggling readers. The study included students in kindergarten and first grade from schools in poor rural counties in Nebraska, New Mexico, North Carolina, and Texas.

The study authors conducted a cluster randomized controlled trial (RCT) in which 16 schools were randomly assigned to conditions: eight to the intervention group and eight to the comparison group. All kindergarten and first-grade classrooms within each school were included in the evaluation. Within each eligible classroom, five struggling and five non-struggling readers were randomly selected for the study. Struggling and non-struggling readers were identified using mandated state assessments, classroom performance, and help from the *TRI* literacy coach during the first 2 months of the school year. One intervention school left the study due to problems with Internet connectivity.

Struggling readers in the intervention group received one-on-one reading instruction from their *TRI*-trained classroom teacher. Non-struggling readers and all students in the comparison condition continued to receive normal reading instruction, without *TRI*, via their regular classroom teacher or other school staff.

The study authors presented two contrasts that tested the effectiveness of the *TRI*.[3] Within the experimental design, the authors' primary contrast compared struggling readers in intervention schools to struggling readers in comparison schools. In the 15 schools that completed the study, there were initially 385 struggling readers (220 in the intervention group and 165 in the comparison group). Between 247 and 250 of these students were included in the analysis, depending on the outcome.[4] The second contrast used a quasi-experimental design (QED) to compare struggling readers in intervention schools with non-struggling readers in the same schools. In addition to examining the impact of *TRI* for kindergarten and first-grade classrooms together, the authors also examined the impact of *TRI* for each grade separately for both the experimental and quasi-experimental contrasts.

The study examined five outcomes within two reading domains—alphabetics and comprehension. The alphabetics outcomes included the Word Attack, Letter Word Identification, and Spelling of Sounds subtests from the Woodcock-Johnson III Diagnostic Reading Battery (WJ-DRB III). The comprehension outcomes included the Peabody Picture Vocabulary Test-III (PPVT-III) and the Passage Comprehension subtest from the WJ-DRB III.

WWC Rating of the Experimental Design

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

The experimental analysis compared struggling readers in randomly assigned intervention and comparison schools. **A subset of the experimental analyses described in the study *meets WWC group design standards with reservations*.** Of the analyses based on these randomly assigned schools, those that combined kindergarten and first-grade students for all five outcomes, as well as the analysis of the Passage Comprehension outcome for the first-grade subgroup, received this rating.

All contrasts of randomly-assigned groups (RCT contrasts) had high attrition. Therefore, the study authors were required to demonstrate equivalence of the intervention and comparison groups to *meet WWC group design standards with reservations*. The intervention/comparison group differences for the sample of struggling readers were less than 0.25 standard deviations for all contrasts. All but one of the baseline differences were between 0.05 and 0.25 standard deviations, but the authors controlled for these differences using pretest scores, student race and gender, and maternal education, demonstrating baseline equivalence for all contrasts.

Among the RCT contrasts conducted separately by grade, study authors demonstrated equivalence for only a single outcome, Passage Comprehension, for the first-grade subgroup. The remainder of the RCT contrasts based on kindergarten and first-grade subgroups *did not meet WWC group design standards* because equivalence of the intervention and comparison groups was required and not demonstrated.

What did the study find?

The study authors found, and the WWC confirmed, that the *TRI* had positive and statistically significant impacts on all three alphabetics outcomes and one of the two comprehension outcomes for struggling readers. The WWC also confirmed that the *TRI* had no statistically significant impact on passage comprehension for the subgroup of first-grade struggling readers.

WWC Rating of the Quasi-Experimental Design

The quasi-experimental research described in this report does not meet WWC group design standards

The quasi-experimental analyses that compared struggling readers in intervention schools to non-struggling readers in the same schools *do not meet WWC group design standards* because baseline equivalence was required and not demonstrated for any contrasts. The remainder of this single study review focuses on the RCT contrasts *that meet WWC group design standards with reservations.*

Features of the *Targeted Reading Intervention (TRI)*

The *TRI* is a professional development intervention designed to support teachers in delivering instruction to struggling readers. All kindergarten and first-grade teachers in schools assigned to receive the *TRI* attended a 3-day summer workshop on *TRI* strategies. During the school year, *TRI* teachers received biweekly observation and feedback from *TRI* literacy coaches. They also met biweekly with other *TRI* teachers and their *TRI* literacy coach to reinforce strategies and problem solve. In addition, *TRI* teachers participated in workshops every few months designed to support understanding of the *TRI* process, models, and strategies. All school-year support was provided via webcam consultation. The *TRI* also includes a program website with instructional resources and the ability to interact with coaches via email.

During the school year, teachers delivered ongoing one-on-one reading instruction for struggling readers in 15-minute sessions. Each session included the following three components: (1) re-reading selected texts for fluency; (2) “word work” using letter tiles to demonstrate the alphabetic principle, teach phoneme–grapheme relationships, support phonemic awareness development, and improve student recognition of sight words; and (3) guided oral reading. Teachers worked individually with each student for an average of 14 sessions over the course of the year. When students made rapid progress, they were placed in a small group, and another struggling reader began receiving one-on-one instruction from the teacher.

Appendix A: Study details

Vernon-Feagans, L., Kainz, K., Hedrick, A., Ginsberg, M., & Amendum, S. (2013). Live webcam coaching to help early elementary classroom teachers provide effective literacy instruction for struggling readers: The Targeted Reading Intervention. *Journal of Educational Psychology, 105*(4), 1175–1187. Retrieved from https://eric.ed.gov/?id=EJ1054424

Setting

The study was conducted in kindergarten and first-grade classrooms in public schools in poor rural counties in Nebraska, New Mexico, North Carolina, and Texas.

Study sample

Sixteen rural schools were assigned to matched pairs based on district, school size, school participation in *Reading First*, and the percentages of students who were minorities and receiving free or reduced-price lunch.[5] One school in each pair was randomly selected for the intervention. One intervention school later left the study due to problems with Internet connectivity. All kindergarten and first-grade classrooms in the remaining 15 schools were included in the study. Teachers identified struggling readers using state assessment data, classroom observation information, and input from the *TRI* literacy coach. Within each classroom, five struggling readers were randomly selected for the study, for a total of 385 students (220 in the intervention group and 165 in the comparison group).[6] The analytic sample included 247–250 students (158–160 in the intervention group and 87–90 in the comparison group, depending on study outcome) across 15 schools (seven intervention and eight comparison). Over half of the sample was male (63% in the intervention group and 54% in the comparison), and less than half were White (49% in the intervention group and 39% in the comparison group).

Intervention group

Teachers in *TRI* schools attended a 3-day summer workshop on *TRI* strategies. During the school year, *TRI* teachers received biweekly observation and feedback from *TRI* literacy coaches, and met biweekly with other *TRI* teachers and their *TRI* literacy coach to reinforce strategies and problem solve. *TRI* teachers also participated in workshops every few months to obtain support with understanding of the *TRI* process, models, and strategies. All school-year support was provided via webcam. The *TRI* program also provides a website with instructional resources and the ability to interact with coaches via email.

During the school year, teachers delivered ongoing one-on-one reading instruction for struggling readers in 15-minute sessions. Each session included the following three components: (1) re-reading selected texts for fluency; (2) "word work" using letter tiles to demonstrate the alphabetic principle, teach phoneme–grapheme relationships, support phonemic awareness development, and improve student recognition of sight words; and (3) guided oral reading (more detail on the three components is provided in the study). Teachers worked individually with a student for an average of 14 sessions over the course of the year. When students made progress, they were placed in a small group, and another struggling reader began receiving one-on-one instruction from the teacher.

Comparison group

Students in the comparison condition continued to receive normal reading instruction, without the use of the *TRI*, via their regular classroom teacher or other school staff.

Outcomes and measurement

To measure program impacts, study authors used outcomes from five standardized tests in two domains—alphabetics and comprehension. The alphabetics outcomes include the Word Attack, Letter Word Identification, and Spelling of Sounds subtests from the Woodcock-Johnson III Diagnostic Reading Battery (WJ-DRB III). The comprehension outcomes include the Peabody Picture Vocabulary Test-III (PPVT-III) and the Passage Comprehension subtest of the WJ-DRB III. The tests were taken in the spring of the school year when the intervention was complete. For more details on these outcome measures, see Appendix B.

Support for implementation

All of the *TRI* coaches had experience as teachers and/or reading coaches in early elementary school. Most were doctoral students in education. The coaches received feedback from the intervention director by providing videotapes of their own teaching of individual students. The coaches received additional feedback throughout the school year, with a particular focus on how to motivate teachers to implement the *TRI* well.

Reason for review

This study was identified for review by the WWC because it was supported by a grant to the University of North Carolina, Chapel Hill (Principal Investigator: Tom Farmer) from the National Research Center for Rural Education Support at the Institute of Education Sciences (IES).

Appendix B: Outcome measures for each domain

Alphabetics	
Woodcock-Johnson III Diagnostic Reading Battery (WJ-DRB III) Letter Word Identification	This is a subtest of the WJ-DRB III that measures how well students can identify words. The items range in difficulty, from requiring students to identify individual letters to requiring students to pronounce words that are not commonly used in written English. The measure has a median reliability of .91 in the 5–19 age range.
WJ-DRB III Spelling of Sounds	This is a subtest of the WJ-DRB III that measures how well students can use phonological and orthographical skills to spell sounds. Easier items require single-letter responses. More difficult items cover letter combinations that can become complex. The measure has a median reliability of .74. The authors did not specify the age range for the reliability.
WJ-DRB III Word Attack	This is a subtest of the WJ-DRB III that measures phonic and structural analysis skills needed to pronounce unfamiliar sounds and words. The items range in difficulty, from requiring students to pronounce single letter sounds to requiring students to pronounce multi-letter sounds that are either non-words or low-frequency words. The measure has a median reliability of .87 in the 5–19 age range.
Comprehension	
Peabody Picture Vocabulary Test-III	This is a standardized test of vocabulary in which students are asked to identify which picture matches the word spoken by an examiner. Alpha coefficients for elementary age students range from .92 to .95 on this test.
WJ-DRB III Passage Comprehension	This is a subtest of the WJ-DRB III that measures passage comprehension. The easier items use pictures to help measure comprehension, while the more difficult items require students to provide missing key words in longer passages. The measure has a median reliability of .83. The authors did not specify the age range for the reliability.

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	
Alphabetics								
Woodcock-Johnson III Diagnostic Reading Battery (WJ-DRB III) Letter Word Identification	Struggling readers	15 schools/ 247 students	nr	nr	nr	0.50	+19	.02
WJ-DRB III Spelling of Sounds	Struggling readers	15 schools/ 248 students	nr	nr	nr	0.39	+15	.02
WJ-DRB III Word Attack	Struggling readers	15 schools/ 249 students	nr	nr	nr	0.39	+15	.04
Domain average for alphabetics						**0.43**	**+16**	**Statistically significant**
Comprehension								
Peabody Picture Vocabulary Test-III	Struggling readers	15 schools/ 247 students	nr	nr	nr	–0.17	–7	.36
WJ-DRB III Passage Comprehension	Struggling readers	15 schools/ 250 students	nr	nr	nr	0.46	+18	.01
Domain average for comprehension						**0.41**	**+6**	**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 individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The 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. nr = not reported.

Study Notes: In response to an author query, the authors provided *d*-statistics and *p*-values from their hierarchical linear model (HLM) using unimputed data. The *d*-statistics (a type of standardized effect size) reported by the authors were calculated by dividing the HLM coefficient by the square root of the total variation in the model, which is similar to the calculation that underlies the WWC's preferred effect size measure. To approximate the WWC's preferred measure, Hedges' *g* corrected for small-sample bias, the WWC multiplied the author-reported *d*-statistic by the WWC small sample size correction. For more information, please refer to the WWC Standards and Procedures Handbook (version 3.0). The *p*-values presented here are those reported by the authors in response to an author query.

A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. This study is characterized as having a statistically significant positive effect for the alphabetics and comprehension domains because the effect for at least one measure within each domain is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Standards and Procedures Handbook (version 3.0), p. 26.

Appendix D: Supplemental findings for the first-grade sample

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	
Comprehension								
WJ-DRB III Passage Comprehension	Struggling readers in first grade	15 schools/ 128 students	462.15 (14.62)	455.74 (17.43)	6.41	0.41	+16	.18

Table Notes: The supplemental findings presented in this table are additional findings that meet WWC design standards with or without reservations, but do not factor into the determination of the study 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 individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding.

Study Notes: In response to an author query, the authors provided the posttest means and standard deviations for the sample of first-grade students with pretest data. The WWC calculated the program group mean using a difference-in-differences approach by adding the impact of the program (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 *p*-value presented here was calculated by the WWC and includes a correction for clustering. No *p*-value was reported for this contrast in the original study.

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 Review of Individual Studies Protocol (version 3.0).

[2] This study was previously reviewed for a grant competition in February 2016, and received a rating of meets WWC group design standards with reservations for two contrasts: (1) the comparison of struggling readers in intervention schools to struggling readers in comparison schools and (2) the comparison of non-struggling readers in intervention schools to non-struggling readers in comparison schools. The current review found that only the first contrast tested the effectiveness of *TRI*. The second contrast did not test the effectiveness of *TRI* because the non-struggling readers in both groups did not receive *TRI*. Therefore, only the first contrast was eligible for review and again received a rating of meets WWC group design standards with reservations. This change also affected the statistical significance of the findings. In the previous review, outcomes from both contrasts were included in the WWC's adjustments for multiple comparisons; after the adjustments, none of the outcomes remained statistically significant. In the current review, only outcomes from the first contrast were included in the WWC's adjustments for multiple comparisons, and four of the five outcomes remained statistically significant after the adjustments.

[3] The study also contained two additional contrasts that did not examine the effect of the intervention: (1) struggling readers in comparison schools compared to non-struggling readers in the same schools and (2) non-struggling readers in intervention schools compared to non-struggling readers in comparison schools. These contrasts do not meet WWC screening criteria because in both cases neither condition received the intervention *(TRI).*

[4] The authors provided sample sizes for the experimental analyses using unimputed data in response to an author query.

[5] *Reading First* is a grant program for states and districts, administered by the U.S. Department of Education, with the goal of using evidence-based instruction to ensure that students read well by the end of the third grade.

[6] The baseline sample sizes were provided by the authors in response to an author query.

Recommended Citation

What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2017, June). *WWC review of the report: Live webcam coaching to help early elementary classroom teachers provide effective literacy instruction for struggling readers: The Targeted Reading Intervention.* Retrieved from https://whatworks.ed.gov

Glossary of Terms

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

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

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

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

Domain A domain is a group of closely related outcomes.

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

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

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

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

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

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

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

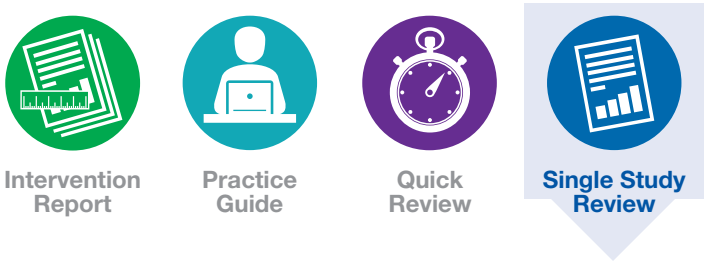
Single-case design (SCD) A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

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

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

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

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



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

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