

WWC Review of the Report “The Effectiveness of a Program to Accelerate Vocabulary Development in Kindergarten”^{1,2}

The findings from this review do not reflect the full body of research evidence on *Kindergarten PAVEd for Success*.

What is this study about?

The study examined whether exposure to *Kindergarten PAVEd for Success*, a vocabulary instruction program, improved the expressive vocabulary of kindergartners.

The study analyzed data for nearly 1,300 kindergarten students in 64 schools serving predominantly rural and high poverty youth in the Mississippi Delta region and surrounding areas.

Eligible schools were placed into three blocks based on previous participation in reading initiatives and then randomly assigned within blocks to either supplement their language arts curriculum with the *Kindergarten PAVEd for Success* program or not. Prior to random assignment of schools, two kindergarten classrooms were randomly selected from each school for participation in the study; a random sample of 10 students was then drawn from each classroom. The study followed this sample of students in each school.

The study assessed the *Kindergarten PAVEd for Success* program’s effectiveness by comparing the expressive vocabulary and listening comprehension of students in the treatment and comparison groups at the end of the school year.

WWC Rating

The research described in this report meets WWC evidence standards without reservations

Strengths: This study is a well-implemented randomized controlled trial.

Features of the *Kindergarten PAVEd for Success* Program

The *Kindergarten PAVEd for Success* program is a 24-week in-class supplement to a school’s core language arts program. It is built around three components:

- (1) *Explicit Vocabulary Instruction* on a large set of target words aligned with themes in Mississippi’s science and social studies frameworks;
- (2) *Interactive Book Reading*, which involves teachers asking questions that promote comprehension and oral language skills during story-reading time; and
- (3) *Adult-Child Conversations*, in which teachers have frequent conversations with individual or small groups of students to introduce or use new vocabulary.

What did the study find?

Kindergarten students in schools using *Kindergarten PAVEd for Success* as a supplement to regular literacy instruction performed better than kindergarten students in comparison schools. The average effect size of 0.12 in the reading comprehension domain was statistically significant.

The authors reported that students who received *Kindergarten PAVEd for Success* instruction were one month ahead in vocabulary development at the end of kindergarten, compared with students in the comparison group.

Appendix A: Study details

Goodson, B., Wolf, A., Bell, S., Turner, H., & Finney, P. B. (2010). *The effectiveness of a program to accelerate vocabulary development in kindergarten (VOCAB)* (NCEE 2010-4014). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Setting Sample schools were selected from the Mississippi Delta region and surrounding districts. Students at these schools tended to be at risk for poor reading outcomes because of high levels of family poverty.

Study sample Eligible schools were defined as those that had at least two kindergarten teachers agreeing to participate and at least 40% of students who were eligible for free and reduced-price lunch. Eligible students were defined as those who:

- were enrolled in kindergarten at the end of the first month of the school year,
- had parental permission to participate in the study,
- were not learning English as a second language, and
- did not have a hearing or language impairment that could prevent them from taking assessment tests.

The baseline study sample consisted of 619 treatment students (in 62 classrooms and 31 schools) and 700 comparison students (in 68 classrooms and 34 schools). The analysis sample was missing one treatment school (with 23 students) that dropped out of the study. Post-test data were collected for 569 treatment students and 659 comparison students.

Intervention group *Kindergarten PAVEd for Success* was adapted for kindergarten from *PAVEd for Success (PAVE)*, an earlier preschool version. The program's overall goal is to introduce around 240 target words in 24 weekly sessions by presenting 10 target words per session. The source of the vocabulary words for this study was the Mississippi social studies and science framework. The words were taught using explicit instruction and were reinforced through repeated exposure to three core techniques: storybook reading, extension activities, and classroom conversations and discussion. Of the 12 target strategies, implementation analyses suggest that 32% of intervention teachers implemented 10 strategies, 37% implemented eight or nine, 25% implemented five to seven, and less than 10% of teachers implemented four or fewer strategies.

Comparison group Schools in the comparison condition implemented their existing language arts curriculum.

Outcomes and measurement Prior to the introduction of the intervention and after its completion, students took the Expressive Vocabulary Test-2 (EVT-2) and the Listening Comprehension subtest of the Kaufman Test of Educational Achievement-II (KTEA-II). For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

For this study, teachers received an initial group training that lasted two days, as well as three follow-up telephone conference calls on implementation issues. Follow-up implementation support was provided in two rounds of site visits. Classroom observations and remediation were conducted by a team from Regional Educational Laboratory Southeast, with oversight by the curriculum developer. Implementation findings suggest that all intervention teachers attended the initial training and received two rounds of fidelity observation, but just 48% participated in all three follow-up conference calls.

Reason for review

This study was identified for review by the WWC because it is an Institute of Education Sciences (IES)-funded study conducted by 2006–11 Regional Educational Laboratory Southeast administered by the SERVE Center at the University of North Carolina at Greensboro.

Appendix B: Outcome measures for each domain

Reading comprehension

Expressive Vocabulary Test–2 (EVT-2)

EVT-2 is a standardized test that is administered one-on-one and measures students' expressive vocabulary. Students are presented with a pictured stimulus and asked to provide a single-word response. For this study, scores were normed using a national sample of children of the same age and were standardized to have a mean of 100 and a standard deviation of 15 points.

Kaufman Test of Educational Achievement–II (KTEA-II), Listening Comprehension subtest

KTEA-II is a normed test that measures students' ability to answer comprehension questions about short passages that have been read aloud by an adult. For this study, scores were normed using a national sample of children of the same age and were standardized to have a mean of 100 and a standard deviation of 15 points.

Appendix C: Study findings by 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 | |
| Reading comprehension | | | | | | | | |
| <i>Expressive Vocabulary Test–2 (EVT-2)</i> | Kindergarten students | 64 schools/ 1,228 students | nr (nr) | nr (nr) | 1.45 | 0.13 | +5 | 0.02 |
| <i>Kaufman Test of Educational Achievement–II (KTEA-II), Listening Comprehension subtest</i> | Kindergarten students | 64 schools/ 1,228 students | nr (nr) | nr (nr) | 1.39 | 0.11 | +4 | 0.14 |
| Domain average for reading comprehension | | | | | | 0.12 | +5 | Statistically significant |

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. 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. 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; the study is characterized as having a statistically significant positive effect because univariate statistical tests are reported for each outcome measure, the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant. nr = not reported.

Study Notes: The study used a three-level hierarchical linear modeling (HLM) analysis, which accounts for clustering at the classroom and school levels. The mean differences, effect sizes, and p-values reported here are based on nonimputed data and were reported in the original study; sample sizes, means, and standard deviations for the nonimputed data were not reported in the study. Missing data were imputed separately for treatment and comparison groups on demographic covariates, pretest scores, and posttest scores, using single stochastic regression. Data were imputed for less than 10% of students in the treatment and comparison groups. Sensitivity analyses indicate that the magnitude, standard errors, and statistical significance of impact estimates were similar regardless of whether missing pretest data, missing posttest data, or missing covariate data were imputed. The sample sizes reported here represent the students for whom posttest data were collected. A correction for multiple comparisons was needed but did not affect significance levels.

Endnotes

¹ Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the author[s]) to assess whether the study design meets WWC evidence standards. The review reports the WWC's assessment of whether the study meets WWC evidence standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. The WWC rating applies only to the summarized results, and not necessarily to all results presented in the study. This study was reviewed using the Beginning Reading review protocol, version 2.0.

² Absence of conflict of interest: The Regional Educational Labs were provided technical assistance by Mathematica Policy Research, which also operates the WWC. For this reason, this study was reviewed by staff from subcontractor organizations.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, June). *WWC review of the report: The effectiveness of a program to accelerate vocabulary development in kindergarten*. 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 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 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 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% ($p < 0.05$). |
| 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 \(version 2.1\)](#) for additional details.