Effective Programs for Struggling Readers: A Best-Evidence Synthesis

Robert E. Slavin
Johns Hopkins University
-and-
University of York

Cynthia Lake
Johns Hopkins University

Susan Davis
Success for All Foundation

Nancy A. Madden
Johns Hopkins University
-and-
University of York

June, 2009

This research was funded by the Institute of Education Sciences, U.S. Department of Education (Grant No. R305A040082). However, any opinions expressed are those of the authors and do not necessarily represent IES positions or policies.

We thank Steve Ross, Alan Cheung, Bette Chambers, Yola Center, and Jean Gross for comments on an earlier draft. We also thank Kate Misso and Michele Victor for assistance in locating references.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Effective Programs for Struggling Readers: A Best-Evidence Synthesis

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Abstract

This article reviews research on the achievement outcomes of alternative approaches for struggling readers in grades K-5: One-to-one tutoring, small group tutorials, classroom instructional process approaches, and computer-assisted instruction. Study inclusion criteria included use of randomized or well-matched control groups, study duration of at least 12 weeks, and use of valid measures independent of treatments. A total of 96 studies met these criteria. The review concludes that one-to-one tutoring is very effective in improving reading performance. Tutoring models that focus on phonics obtain much better outcomes than others. Teachers are more effective than paraprofessionals and volunteers as tutors. Small-group, phonetic tutorials can be effective, but are not as effective as one-to-one phonetically-focused tutoring. Classroom instructional process programs, especially cooperative learning, can have very positive effects for struggling readers. Computer-assisted instruction generally had few effects on reading. Taken together, the findings support a strong focus on improving classroom instruction and then providing one-to-one, phonetic tutoring to students who continue to experience difficulties.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Over the past 25 years, there have been extraordinary developments in research, policy, and practice relating to programs for elementary-aged children who are struggling to learn to read. While there has long been concern about reading disabilities, dyslexia, and underachievement, research and development since the 1980's has created a sense of optimism that most children who start off their time in school struggling to learn to read can be quickly brought into the mainstream in this crucial skill. The appearance of *Reading Recovery*, first in New Zealand and later in the U.S., the U.K., and throughout the English-speaking world, gave particular hope that tutors with extensive training could prevent reading failure with a substantial proportion of the children who were failing in first grade and were therefore at risk of serious difficulties throughout their time in school. In the 1990's, the Clinton administration's America Reads initiative encouraged the creation of programs for volunteer tutors to work with struggling children, and this led to widespread development and evaluation of replicable programs for this purpose. Reading First, the Bush Administration's initiative for children in grades K-3, focused on high-poverty, low-achieving schools, with a particular focus on small-group interventions for struggling readers. A new focus in special education on response to intervention (see Allington & Walmsley, 2007, Fuchs & Fuchs, 2006; Gersten et al., 2009), in which at risk children receive small-group interventions and then possibly one-to-one tutoring to attempt to solve their problems before they might be referred to special education, has also encouraged development and research on small group tutorials. In the U.K., the Labour government introduced a program called Every Child a Reader to disseminate *Reading Recovery* throughout England (see

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Burroughs-Lange, 2007, 2008; Policy Exchange, 2009). In addition to benefitting children, each of these initiatives has stimulated research of all kinds. In 1993, Wasik & Slavin reviewed research on tutoring programs, and found only five studies. In 2000, Elbaum, Vaughn, Hughes, & Moody reviewed one-to-one tutoring programs, but the great majority of the available research focused on just one program, *Reading Recovery*. Today, there are many programs designed to help struggling readers, and there is much research on factors that affect the impact of tutoring and other interventions. There is also much research on the effects on struggling readers of classroom programs and comprehensive school reform models that impact entire schools.

The importance of getting children off to a good start in reading cannot be overstated. In the elementary grades, success in school is virtually synonymous with success in reading, and children without strong reading skills by middle school are headed for disaster. Children who fail to read in the early grades incur so many costs to the education system, in special education, remediation, grade repetition, delinquency, and ultimate dropout, that even very expensive interventions can be justified on cost-effectiveness grounds alone, while at the same time preventing damage to young peoples' lives. Further, reading failure is not distributed randomly, but is concentrated among schools serving many disadvantaged, minority, and limited English proficient children. It is in the early elementary grades where the gap in performance between children of different races first appears, and this gap is perhaps the most important policy issue in education in the U.S. On the fourth grade National Assessment of Educational Progress (NAEP, 2007), 43% of White children achieved at the "proficient" level on the National Assessment of

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Educational Progress, but only 14% of African American, 17% of Hispanic, and 8% of American Indian children scored at this level. Effective reading programs are important for children of all backgrounds, but for disadvantaged and minority children and for children with learning disabilities, who particularly depend on school to achieve success, effective reading programs are especially important.

Because of the importance of ensuring success in reading for all children, the policy focus on the problem, and the costs involved, it is especially important to understand which types of programs are most likely to have a strong and lasting impact on the reading success of struggling children. The purpose of the present article is to review research on all types of approaches that have been evaluated as solutions for struggling readers. As noted earlier, portions of this research have been reviewed previously with conflicting conclusions. Wasik & Slavin (1994) reviewed research on one-to-one tutoring, concluding that tutoring had substantial positive impacts and that certified teachers obtained better results than paraprofessionals. Wasik (1997) then found positive outcomes for a variety of volunteer tutoring programs, such as those used in *America Reads*. Elbaum et al. (2000), in a review of one-to-one tutoring, focused primarily on *Reading Recovery,* finding positive effects but concluding (based on two small dissertations) that small group tutorials might be as effective as one-to-one tutoring. Shanahan & Barr (1995) reviewed research on *Reading Recovery* and while recognizing its effectiveness, questioned its cost-effectiveness. In a recent What Works Clearinghouse practice guide, Gersten et al. (2009), in providing a research case for response to intervention, concluded that there is

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

strong evidence to support use of small-group tutorials with struggling readers, based on 11 studies that fully or partially met the standards of the What Works Clearinghouse (WWC). The WWC (2009) Beginning Reading Topic Report gave top ratings to *Reading Recovery*, though not to any of the programs cited by Gersten et al. (2009) as evidence for the effectiveness of small group instruction.

Few researchers or educators would doubt that one-to-one tutoring is effective for struggling readers. The questions today are beyond this and are critical for providing useful guidance to educators. After 25 years of research and reform, we want to know about long-term impacts of early intervention, about different effects of tutoring by certified teachers as opposed to paraprofessionals and volunteers, and whether small group interventions can be as effective as one-to-one tutoring. We want to know the importance of a focus on phonics and phonological awareness in reading interventions. We want to know how to build on success in the early grades to maintain gains into secondary schools, and we want to know how improved classroom programs and technology might contribute to the success of struggling readers. Now that there is a rich diversity of approaches that have been rigorously evaluated, it is no longer enough to review tutoring or small group interventions in isolation. Educators and policy makers deserve comparative, fair, and readily understood information about the likely impacts of alternative approaches or combinations of approaches. We also want to know more about for whom various models are likely to work and under what conditions. These are the questions that drive the current review.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

The present article reviews research on the achievement outcomes of practical approaches for struggling readers, applying consistent methodological standards to the research. It is intended to provide fair summaries of the achievement effects of the full range of remedial and preventive reading approaches available to educators and policy makers, and to summarize for researchers the current state of the art in this area. The scope of the review includes all types of programs that teachers, principals, or superintendents might consider to improve the success of their struggling readers: Tutoring, small group tutorials, teacher professional development for effective core reading instruction, and instructional technology. The review uses a form of best evidence synthesis (Slavin, 1986), adapted for use in reviewing "what works" literatures in which there are generally few studies evaluating each of many programs (see Slavin, 2008). It is part of a series, all of which used the same methods with minor adaptations. Separate research syntheses review research on beginning and upper-elementary reading programs (Slavin, Lake, Chambers, Cheung, & Davis, in press), middle and high school reading programs (Slavin, Cheung, Groff, & Lake, 2008), and reading programs for English language learners (Cheung & Slavin, 2005).

The synthesis of research on beginning reading programs (Slavin et al., in press a) provides the closest background for the present review. The beginning reading review identified 63 studies that met the inclusion standards. These were divided into four categories: reading curricula (core and supplementary textbooks), instructional technology, instructional process programs (such as cooperative learning), and combinations of curricula and instructional process.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Effect sizes for curricula (ES=+0.12) and for instructional technology (ES=+0.11) were low. Larger effect sizes (ES=+0.31) were found for instructional process programs, especially cooperative learning programs in which students help one another master reading comprehension skills in small teams or pairs.

Focus of the Current Review

The present review uses procedures similar to those used in the beginning reading review to examine research on programs for struggling readers. The purpose of the review is to place all types of interventions intended to enhance the achievement of students experiencing difficulties in learning to read on a common scale, to provide educators and policy makers with meaningful, unbiased information that they can use to select approaches most likely to make a difference with their students. The review emphasizes practical programs that are or could be used at scale. It therefore emphasizes large studies done over significant time periods using standard measures to maximize the usefulness of the review to educators. The review also seeks to identify common characteristics of programs likely to make a difference in reading achievement. This synthesis was intended to include all kinds of approaches to reading interventions for struggling readers' instruction, and groups them in six categories as follows:

1. One-to-One Tutoring by Teachers. This includes *Reading Recovery* and other tutoring models.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

2. One-to-One Tutoring by Paraprofessionals and Volunteers
3. Small Group Tutorials
4. Classroom Instructional Process Approaches. This category includes cooperative learning (e.g., *Cooperative Integrated Reading and Composition (CIRC)* and *Peer-Assisted Learning Strategies (PALS)), Direct Instruction,* and other approaches that focus on training teachers in effective classroom methods rather than pullout services for at-risk students. These studies report outcomes separately for students in the lowest performance levels of their grades, even though they affect all children.
5. Classroom Instructional Process with Tutoring. This category consists of studies of *Success for All*, which trains teachers in reading strategies, cooperative learning, and other methods throughout the elementary grades, and provides 1-1 tutoring by teachers to low achieving children. Outcomes are reported for students in the lowest performance levels of their grades.
6. Instructional Technology. These studies report effects of computer assisted instruction and other uses of technology for children in the lowest performance levels of their classes.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Methodological Issues Characteristic of Research on Struggling Readers

While a review of research on reading programs for struggling readers shares methodological issues common to all systematic reviews, there are also some key issues unique to this topic. One of these relates to measurement. In the early stages of reading, researchers often use measures such as phonemic awareness that are not "reading" in any sense, though they are precursors. However, measures of reading comprehension and reading vocabulary tend to have floor effects at the kindergarten and first grade levels. The present review accepted measures such as letter-word identification and word attack as measures of reading, but did not accept measures such as auditory phonemic awareness. Measures of oral vocabulary, spelling, and language arts were excluded.

Another problem of early reading measurement is that in kindergarten studies, it is possible to find positive effects of programs that introduce skills not ordinarily taught in kindergarten on measures of those skills. For example, until the late 1990's it was not common in U.S. kindergartens for children to be taught phonics or phonemic awareness. Programs that moved these then first-grade skills into kindergarten might appear very effective in comparison to control classes receiving little or no instruction on them, but would in fact simply be teaching skills the children would probably have mastered somewhat later. For this reason, kindergarten interventions are included in this review only if they followed students to the end of first grade or later, by which time it is certain that control students as well as experimental students would have been formally taught to read.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Review Methods

As noted earlier, the review methods used here are similar to those used by Slavin, Lake, Chambers, Cheung, & Davis (in press), who adapted a technique called best-evidence synthesis (Slavin, 1986). Best-evidence syntheses seek to apply consistent, well-justified standards to identify unbiased, meaningful information from experimental studies, discussing each study in some detail, and pooling effect sizes across studies in substantively justified categories. The method is very similar to meta-analysis (Cooper, 1998; Lipsey & Wilson, 2001), adding an emphasis on narrative description of each study's contribution. It is similar to the methods used by the What Works Clearinghouse (2009), with a few important exceptions noted in the following sections. See Slavin (2008) for an extended discussion and rationale for the procedures used in this series of best-evidence reviews.

Literature Search Procedures

A broad literature search was carried out in an attempt to locate every study that could possibly meet the inclusion requirements. Electronic searches were made of educational databases (JSTOR, ERIC, EBSCO, Psych INFO, Dissertation Abstracts) using different combinations of key words (for example, "elementary or primary students," "reading," "achievement") and the years 1970-2009. Results were then narrowed by subject area (for example, "reading intervention," "educational software," "academic achievement," "instructional strategies"). In addition to looking for studies by key terms and subject area, we conducted

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

searches by program name. Web-based repositories and education publishers' websites were also examined. We attempted to contact producers and developers of reading programs to check whether they knew of studies that we had missed. Citations were obtained from other reviews of reading programs including the What Works Clearinghouse (2009) beginning reading topic report, Gersten et al. (2009), Ritter, Barnett, Denny, & Albin (2009), Elbaum et al. (2000), Wasik & Slavin (1994), Wasik (1997), National Reading Panel (2000), Snow, Burns & Griffin (1998), Torgerson, Brooks, & Hall (2006), and Brooks (2007). We also conducted searches of recent tables of contents of key journals. We searched the following tables of contents from 2000 to 2008: *American Educational Research Journal, Reading Research Quarterly, Journal of Educational Research, Journal of Educational Psychology, Reading and Writing Quarterly, British Educational Research Journal,* and *Learning and Instruction.* Citations of studies appearing in the studies found in the first wave were also followed up.

Effect Sizes

In general, effect sizes were computed as the difference between experimental and control individual student posttest means after adjustment for pretests and other covariates, divided by the unadjusted posttest control group standard deviation. If the control group SD was not available, a pooled SD was used. Procedures described by Lipsey & Wilson (2001) and Sedlmeier & Gigerenzor (1989) were used to estimate effect sizes when unadjusted standard deviations were not available, as when the only standard deviation presented was already

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

adjusted for covariates or when only gain score SD's were available. If pretest and posttest means and SD's were presented but adjusted means were not, effect sizes for pretests were subtracted from effect sizes for posttests. In multiyear studies, effect sizes were usually reported for each year but only the final year of treatment is presented in the tables. However, if there were multiple cohorts (e.g., K-1, K-2, K-3), each with adequate pretests, all cohorts were included in the tables.

Effect sizes were pooled across studies for each program and for various categories of programs. This pooling used means weighted by the final sample sizes. The reason for using weighted means is to maximize the importance of large studies, as the previous reviews and many others have found that small studies tend to overstate effect sizes (see Rothstein et al., 2005; Slavin, 2008; Slavin & Smith, in press).

Statistical significance is reported in this review as it was in the original articles, but is not reported if the article did not report significance for a given comparison. For example, if a study reported outcomes separately for high, average, and low achievers and reported treatment by achievement level interactions, but did not report a separate analysis for low achievers, we would report the effect size for low achievers but not statistical significance.

Criteria for Inclusion

Criteria for inclusion of studies in this review were as follows.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

1. The studies evaluated specific, potentially replicable programs for children who are having difficulties learning to read in grades K-5. These are defined as children with reading disabilities, children in the lowest 33% (or lower) of their classes, or any children receiving tutoring or other intensive services to prevent or remediate serious reading problems. Studies of variables (e.g., mainstreaming or after school attendance) or of government funding streams (e.g., Title I, Reading First, Supplemental Educational Services) are not included.
2. The included studies compared children taught using a given reading program to those in a control group taught using an alternative program or standard methods. Studies that compared two experimental methods without a control group representing ordinary practice are not included in the main tables but are discussed as appropriate to answer relevant questions.
3. Studies could have taken place in any country, but the report had to be available in English.
4. Random assignment or matching with appropriate adjustments for any pretest differences (e.g., analyses of covariance) had to be used. Studies without control groups, such as pre-post comparisons and comparisons to "expected" scores, were excluded.
5. Pretest data had to be provided. Studies with pretest differences of more than 50% of a standard deviation were excluded because, even with analyses of covariance, large pretest

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

differences cannot be adequately controlled for, as underlying distributions may be fundamentally different (Shadish, Cook, & Campbell, 2002).

6. The dependent measures included quantitative measures of reading performance, such as standardized reading measures. Experimenter-made measures were accepted if they were comprehensive measures of reading, which would be fair to the control groups, but measures of reading objectives inherent to the experimental program (but unlikely to be emphasized in control groups) were excluded. Studies using measures inherent to treatments, usually made by the experimenter or program developer, have been found to be associated with much larger effect sizes than are measures that are independent of treatments (Slavin & Madden, in press), and for this reason, effect sizes from treatment-inherent measures were excluded. The exclusion of measures inherent to the experimental treatment is a key difference between the procedures used in the present review and those used by the What Works Clearinghouse. As noted earlier, measures of pre-reading skills such as phonological awareness, as well as reading-related outcomes such as oral vocabulary, language arts, and spelling, were not included in this review.
7. A minimum study duration of 12 weeks was required. This requirement is intended to focus the review on practical programs intended for use for the whole year, rather than brief investigations. Study duration is measured from the beginning of the treatments to posttest, so, for example, an intensive 8-week intervention in the fall of first grade would be considered a year-long study if the posttest were given in May. The 12-week criterion

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

has been consistently used in all of the systematic reviews done previously by the current authors. This is another difference between the current review and the What Works Clearinghouse (2009) beginning reading topic report, which included very brief studies.

8. Studies had to have at least 15 students and two teachers or tutors in each treatment group.

Appendix 1 lists studies that were considered germane but were excluded according to these criteria, as well as the reasons for exclusion.

Limitations

It is important to note several limitations of the current review. First, the review focuses on experimental studies using quantitative measures of reading. There is much to be learned from qualitative and correlational research that can add depth and insight to understanding the effects of reading programs, but this research is not reviewed here. Second, the review focuses on replicable programs used in realistic school settings expected to have an impact over periods of at least 12 weeks. This emphasis is consistent with the review's purpose of providing educators with useful information about the strength of evidence supporting various practical programs, but it does not attend to shorter, more theoretically-driven studies that may also provide useful information, especially to researchers. Finally, the review focuses on traditional measures of reading performance, primarily individually-administered or group-administered

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

standardized tests. These are useful in assessing the practical outcomes of various programs and are fair to control as well as experimental teachers, who are equally likely to be trying to help their students do well on these assessments. The review does not report on experimenter-made measures of content taught in the experimental group but not the control group, even though results on such measures may also be of importance to researchers or educators.

Categories of Research Design

Three categories of research designs were identified. *Randomized experiments* (R) were those in which students, classes, or schools were randomly assigned to treatments, and data analyses were at the level of random assignment. When schools or classes were randomly assigned but there were too few schools or classes to justify analysis at the level of random assignment, the study was categorized as a *randomized quasi-experiment* (RQE) (Slavin, 2008). *Matched* (M) studies were ones in which experimental and control groups were matched on key variables at pretest, before posttests were known. Studies using fully randomized designs (R) are preferable to randomized quasi-experiments (RQE), but all randomized experiments are less subject to bias than matched studies. In the text and in tables, studies of each type of program are listed in this order (R, RQE, M). Within these categories, studies with larger sample sizes are listed first. Therefore, studies discussed earlier in each section should be given greater weight than those listed later, all other things being equal.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

One-to-One Tutoring by Teachers

One-to-one instruction from certified teachers and reading specialists is the gold standard among interventions for struggling readers. It is the most expensive solution, but the expense is more than justified if it can make a substantial difference for children at a critical point in their reading development and therefore reduce later needs for special education, remediation, or grade retention. Studies of one-to-one tutoring by teachers are summarized in Table 1, and then described in the following sections.

==========

Table 1 Here

==========

Reading Recovery

Reading Recovery is by far the most widely researched and widely used tutoring program in the world. Originally developed in New Zealand by Marie Clay (1985), *Reading Recovery* provides extensive training, observation, and feedback to certified teachers, who provide daily 30-minute lessons to the lowest 20-30% of first graders in a school until they are reading at the level of average first graders in their school. In general, teachers work with about four children during half of each school day and teach a regular first grade class the other half of the day. A *Reading Recovery* session involves a) re-reading of a familiar book, b) independent reading of a text at the child's level, c) teaching of letter knowledge, d) composing and writing a sentence, e)

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

re-constructing a cut-up sentence, and f) introducing a new book. The books are leveled readers with predictable text. Over the years, *Reading Recovery* has added more of an emphasis on phonics and decoding skills. Teacher training for *Reading Recovery* involves about 75 contact hours and includes live observations through a one-way glass screen and feedback from expert teacher leaders. The training takes place over an entire school year concurrent with practice with children.

From early on, *Reading Recovery* leaders have placed a strong emphasis on program evaluation, and there are many studies that have evaluated program outcomes. However, there are characteristic elements of many *Reading Recovery* studies that are important to understand. First, most *Reading Recovery* studies use posttest measures from Clay's (1985) Diagnostic Observation Survey. Given particular emphasis is a measure called Text Reading Level, in which children are asked to read aloud from leveled readers, while testers (usually other *Reading Recovery* teachers) record accuracy using a running record. Unfortunately, this and other Diagnostic Observation Survey measures are closely aligned to skills taught in *Reading Recovery* and are considered inherent to the treatment; empirically, effect sizes on these measures are typically much greater than those on treatment-independent measures. For example, the review by Elbaum et al. (2000) reported a mean effect size of +0.64 for Text Reading Level and +0.30 for other oral reading measures. In accord with this review's requirement of treatment-independent measures, Diagnostic Observation Survey measures in studies of *Reading Recovery* are not reported in this review.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

A second typical procedure in *Reading Recovery* evaluations is to divide children into "discontinued" and "not discontinued" categories. Discontinued children are those who met criteria for ending tutoring because they were reading at the level of their schoolmates. "Not discontinued" children are those who did not meet criteria, and are in general referred for special education or other services. A third category is children who received some tutoring but not the 60 sessions held to be a full treatment. Some studies of *Reading Recovery* only include the discontinued (i.e., successful) children. In this review, studies are included only if all children assigned to *Reading Recovery,* discontinued or not, are included in the analyses.

Finally, some *Reading Recovery* studies have the tutors themselves administer the tests used as the outcome measures in the evaluation. An example is a widely cited study by Schwartz (2005). This introduces a strong possibility of bias, and such studies were excluded.

The most important evaluation of *Reading Recovery* was a randomized evaluation in 10 Ohio school districts by Pinnell et al. (1994). The study compared *Reading Recovery* and three variations to control groups. One variation, called *Reading Success*, was identical to *Reading Recovery* except that it provided condensed (but still extensive) training for the teachers. These teachers were new to *RR* while those in the *Reading Recovery* condition had at least two years of experience with the program, so teacher experience is confounded with the differences between the treatments. Another treatment, *Direct Instruction Skills Plan (DISP),* provided one-to-one tutoring but used an alternative to *RR* that emphasized direct instruction, mastery, sequential steps, and a curriculum closely linked to classroom instruction. Finally, *Reading and Writing*

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Group used *Reading Recovery*-trained teachers and *RR*-like procedures and materials in small groups. Comparison groups provided the schools' typical Title I services for struggling first graders, typically remedial pull-out instruction in small groups.

There were a total of 283 children in the study (n=193 for the comparison of full *Reading Recovery* (n=31) to control (n=162), which is what appears in Table 1). Approximately 65% of children received free or reduced-price lunches, 74% were White, and 26% were African American. There were six urban, two suburban, and two rural districts throughout Ohio.

In each district, four relatively high-poverty schools were identified for the study. The schools were randomly assigned to use one of the four treatments and then the 10 lowest-achieving children within each school were randomly assigned either to the school's treatment or to the control group. The three tutoring conditions operated from September to February, while the small group treatment was used all year.

The first testing took place in February. Controlling for pretests, and using a conservative HLM analysis, only the full *RR* group significantly exceeded its control group on Woodcock Reading Mastery (ES=+0.49, $p<.05$) and on Gates MacGinitie (ES=+0.51, $p<.05$). Woodcock and Gates effect sizes were, respectively, +0.04 and +0.27 for *Reading Success*, +0.25 and +0.14 for *DISP*, and +0.23 and +0.23 for *Reading and Writing Group*. However, the February test was only an interim measure, as the *Reading and Writing Group* and control treatments were still in operation through the spring semester. The May testing was considered the main evaluation. Only Gates MacGinitie tests were given. None of the effects were significant. Effect sizes were

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

+0.19 for full *RR*, -0.14 for *Reading Success*, -0.05 for *DISP,* and +0.34 for *Reading and Writing Group*. A September followup assessment only used two scales from the Diagnostic Observation Survey, which did not qualify for this review, but it is interesting to note that on the *Reading Recovery* Text Reading Level measure there were positive followup effects for *Reading Recovery* and (non-significantly) for *Reading and Writing Group,* but no positive effects for the other two tutoring models.

In 1984-85, Marie Clay first introduced *Reading Recovery* to the U.S., training a group of teachers and teacher leaders at Ohio State University. Beginning the following school year, *Reading Recovery* was implemented in 12 Columbus schools, with 32 teachers. A longitudinal evaluation of this implementation was carried out by Pinnell, DeFord, and Lyons (1988). In it, 187 first graders in the lowest 20% of their schools were randomly assigned to *Reading Recovery* (n=126) or to an alternative compensatory program (n=26). At the end of first grade, adjusting for pretests, effect sizes on the CTBS were +0.55 for Comprehension, +0.48 for Vocabulary, and +0.52 for Total Reading. Although *Reading Recovery'*s Text Reading Level measure does not meet the standards of this review, it is interesting to note that a longitudinal followup found effect sizes of +0.80 in first grade, +0.47 in second grade, and +0.26 in third grade.

A large matched evaluation of *Reading Recovery* was carried out by the Austin (Texas) Independent School District (Curry, Griffith, & Williams, 1995). It compared 268 first graders who received *Reading Recovery* tutoring to 285 children in matched comparison schools that did not have *Reading Recovery* programs. The students were 47% African American, 47% Hispanic,

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

and 6% Anglo, and 93% low income. Combining across discontinued students (n=83), not discontinued (n=71), and students who received fewer than 60 lessons (n=114), effect sizes on ITBS adjusting for Metropolitan Readiness Test pretests were -0.16 (n.s.). Excluding the students who received fewer than 60 lessons, the adjusted effect size was -0.04. In both cases, substantial positive gains for children who were discontinued were balanced by low gains among not discontinued children and children who received less than 60 lessons.

A large matched study of *Reading Recovery* was carried out in London by Burroughs-Lange (2008; see also Burroughs-Lange & Douetil, 2007; Burroughs-Lange, 2007). Because this is by far the most recent of the qualifying studies, this implementation of *Reading Recovery* contains more of a phonetic emphasis than the program did in the evaluations of the 1980s and 90s, according to the author (Burroughs-Lange, personal communication, May 3, 2009). This study compared children in 21 matched pairs of high-poverty primary schools. The schools were described as being among the lowest achieving in England, with high proportions of English language learners and children receiving free meals. Children were followed over a 2-year period, from fall of Year 1 to spring of Year 2. Three overlapping groups of children were followed. One consisted of all children in the 21 *Reading Recovery* schools (n=457) and those in the 21 control schools (n=398). However, only a subset of these children received *Reading Recovery* tutoring.

Within the *Reading Recovery* schools there were lowest-achieving children who did not receive *Reading Recovery* tutoring. By the end of Year 1, the *RR* sample could be divided into

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

children who received any amount of tutoring (n=87) and those who did not (n=58). The tutored children were well matched with controls (n=147), so this review focuses on the comparison between tutored students and students in schools in which tutoring was not available.

Children were individually pre- and posttested on several measures. Unfortunately, the testers were experienced *Reading Recovery* teachers, and they were aware of which schools were in the experimental and control groups. This is a serious flaw in the design, as the testers may have been hoping for positive *Reading Recovery* outcomes, and means that findings must be interpreted cautiously. Scales from Clay's Observation Survey were excluded as being inherent to the treatment, but treatment-independent measures qualified for the review.

At the end of Year 1 (Burroughs-Lange, 2007), tutored low achievers in the *Reading Recovery* schools scored substantially higher than control students, adjusting for pretests, on the British Abilities Scale (BAS)-Word Reading II (ES=+0.87, $p<.05$) and on Word Reading and Phonics Skills (WRAPS) (ES=+0.65, $p<.05$). At the end of Year 2 (followup), pretest-adjusted differences were ES=+0.88 on Progress in English 7, an individually-administered test focusing on comprehension, spelling, and grammar and were +0.61 on BAS Word Reading and +0.79 on WRAPS, for a mean of +0.76. On National Curriculum tests, given by the students' own teachers rather than testers, 57.4% of control low achievers scored Level 2 or better, while 83% of *Reading Recovery* low achievers scored at this level.

A criticism of *Reading Recovery* made by Hiebert (1994) is that the program touches so few children that it has little impact on entire age cohorts of children. To test this assertion, the

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

London study assessed outcomes for all children who were in the *Reading Recovery* and control schools in Years 1 and 2, and over the full two-year period, effect sizes were +0.23 ($p<.001$) for WRAPS and +0.40 ($p<.001$) for Progress in English 7.

Hurry & Sylva (2007) carried out a long-term evaluation of *Reading Recovery* and of a phonological training intervention with six year olds in primary schools in England. The tutoring was provided during the 1992-1993 school year. A total of 22 schools that already had a trained *Reading Recovery* tutor were each matched with two similar schools, which were then randomly assigned to phonological training or control conditions. In all schools, the lowest-achieving six children were identified based on scores on the Clay Diagnostic Survey. In *Reading Recovery* schools, the four lowest achieving children received tutoring while the two higher-achieving children served as within-school controls. Pretest differences between the two within-school samples were too large to meet inclusion criteria. However, pretest differences on between-school comparisons between the four children tutored in each *Reading Recovery* school (n=89) and the six lowest children in matched schools (n=109) were acceptable, so only the between-school comparisons are reviewed here. Children were relatively impoverished, with 42% receiving free lunches and 16% English language learners.

Children in *Reading Recovery* were tutored 30 minutes daily, receiving an average of 21 weeks of instruction in 77 sessions. 89% were "discontinued" by *Reading Recovery* standards, but all students were kept in the analysis. Children in control classes received standard methods, which included specialized help with reading averaging 21 minutes per week.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

All students were pretested at the beginning of Year 2 (in fall, 1992) and then posttested in spring of Year 2, spring of Year 3, and fall of Year 6. Posttests were adjusted for a battery of pretests. Effect sizes in comparison to controls were +0.84 ($p<.001$) on the British Ability Scale (BAS) Word Reading and +0.85 ($p<.001$) on Neale Prose Reading at the end of Year 2.

At Year 3 follow-up, effect sizes were +0.41 ($p<.001$) for Word Reading, +0.42 ($p<.001$) for Prose Reading, and +0.38 ($p<.01$) for a measure of Non-Word Reading. However, in fall of Year 6, differences were small and non-significant on the NFER-Nelson Group Reading Test (ES=+0.15, n.s.). It is interesting to note that while immediate effects of phonological awareness training were non-significant at the end of Year 2, they approached the effects of *Reading Recovery* at the end of Year 3 (mean ES=+0.27) and somewhat exceeded *Reading Recovery* in Year 6 (ES=+0.21, n.s.). An exploratory analysis found that *Reading Recovery* had small positive reading effects at Year 6 follow-up for children who had been non-readers at pretest but had no effects for those who had shown some reading skills at pretest.

An evaluation of *Reading Recovery* was carried out in urban schools in New South Wales, Australia by Center, Wheldall, Freeman, Outhred, & McNaughton (1995). In ten schools that had already implemented *Reading Recovery*, Year 1 (six year old) children who scored lowest on the Clay Diagnostic Survey were randomly assigned to *Reading Recovery* or control conditions. Despite random assignment, however, there were substantial pretest differences on several measures, so the comparison did not meet the standards of this review. Fortunately, a matched sample of children in 5 schools that had not implemented *Reading Recovery* was also

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

identified, and the children in this sample did match well with the *Reading Recovery* children. Children were pretested before tutoring began and then posttested 15 weeks later (immediate posttest), 30 weeks later (short-term maintenance), and 12 months after posttest (medium-term maintenance). N's were 22 for *Reading Recovery*, 34 for comparison. Adjusting for pretests, immediate posttest differences were substantial, with effect sizes of +1.55 for the Neale Analysis of Reading Ability, +1.20 for a Passage Reading Test, +0.69 for a Cloze Test of Syntactic Awareness, and +1.33 for a Word Attack Skills Test, for a mean of +1.19. At 15-week follow-up (at the end of Year 1), differences remained strong, with adjusted effect sizes of +1.15 for the Neale, +1.00 for Passage Reading, +0.46 for Cloze, and +0.82 for Word Attack, for a mean of +0.86. These end of year outcomes are shown in Table 1, rather than the immediate outcomes that did not allow the control group to complete its instruction. At the end of Year 2, a year later, effect sizes were +1.04 for the Neale, +1.00 for Passage Reading, +0.96 for Word Attack, and +0.87 for Woodcock Passage Comprehension, for a mean of +0.97.

Escamilla (1994) carried out a matched evaluation of *Reading Recovery/Descubriendo la Lectura* used to teach Spanish-dominant first graders struggling in Spanish reading. Low-achieving children in four Southern Arizona schools using a Spanish adaptation of *Reading Recovery/ Descubriendo la Lectura* (n=23) were compared to similar low-achieving children in two control schools (n=23). Children were pre- and posttested on the Aprenda, a standardized Spanish reading test. *Reading Recovery/ Descubriendo la Lectura* children increased from the 28th to the 41st percentile, while controls were virtually unchanged (ES=+0.30).

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

A study in Detroit by Huggins (1999) followed *Reading Recovery* and matched control students from first to fourth grade. Children from 30 Title I schools who received *Reading Recovery* in first grade (n=70) were matched with students in the same classes who were the lowest achieving students not to receive tutoring (n=52). The two groups did not differ on CAT pretest scores from spring of kindergarten. Adjusting for pretests, there were no differences at the end of first grade on CAT Reading Comprehension (ES=+0.03, n.s.), Reading Vocabulary (ES=-0.15, n.s.), or Total Reading (ES=-0.09, n.s.). On MAT Total Reading tests there were no differences at the end of second grade (ES=-0.17, n.s.), third grade (ES=-0.29, n.s.), fourth grade (ES=-0.24, n.s.), or fifth grade (ES=+0.13, n.s.). On the state's MEAP tests, 38.2% of *Reading Recovery* students and 53.7% of control students scored "satisfactory," for an effect size of -0.32.

Adding Phonics Instruction to Reading Recovery

Iverson and Tunmer (1993), two New Zealand educators, carried out a study in Rhode Island that compared a standard form of *Reading Recovery* (as it was at the time) to a modified form that added significant direct instruction in phonics. A total of 23 schools and 26 *Reading Recovery* teachers were involved, teaching 64 first graders from 34 classrooms (n=32E, 32C). Individual children were matched on pretests. There was also an untutored control group, but this comparison did not meet inclusion standards (posttesting was not done at the same time).

Children were pretested on a Dolch Word Recognition test in October, at discontinuation, and at the end of the school year. The children whose tutors added phonics teaching scored non-

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

significantly higher, adjusting for pretests, on the end-of-year Dolch Word Recognition (ES=+0.23, n.s.). Because there was no qualifying control group representing ordinary practice, this study does not appear in Table 1.

Other One-to-One Tutoring by Teachers

Auditory Discrimination in Depth

Auditory Discrimination in Depth (Lindamood & Lindamood, 1984) is a one-to-one tutoring approach with a strong emphasis on phonological awareness, decoding practice, and reading of books with phonetically controlled vocabulary. Torgesen, Wagner, & Rashotte (1997) carried out a 2 ½ year evaluation of *ADD* with children identified in the first semester of kindergarten as having deficiencies in letter-name knowledge and phonological awareness. Children were 50% White and 49% African American. They were randomly assigned to one of four groups: An adaptation of *ADD* (n=33), tutoring using embedded phonics (n=36), tutoring using regular classroom support aligned with classroom instruction (n=37), or control, with no tutoring (n=32). Teachers reported their regular classroom instructional methods to be most consistent with whole language. Children in all tutoring conditions received 80 minutes of 1-1 supplemental instruction each week over a 2 ½ year period. Teachers gave half of the tutoring to each child and aides gave the other half. Over the 2 ½ year period, children received about 88 hours of tutoring, 47 hours from teachers and 41 hours from aides. Nine teachers were also randomly assigned to conditions within 13 schools. They received 20 hours of initial training and

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

then attended bi-weekly 3-hour inservice meetings throughout the study. Aides received 6 hours of training and met once a month for ongoing inservice.

Children were followed into the end of second grade if they were promoted each year, or to the end of first grade if not. Adjusting for pretests (Verbal IQ and letter-word knowledge), effect sizes for the full *ADD* model compared to control were +0.65 for Woodcock Word Identification, +1.02 for Woodcock Word Attack, +0.39 for Woodcock Passage Comprehension, +1.28 for Word Efficiency, and +1.17 for Non-word Efficiency. The authors note the statistically significant and substantial differences in outcomes for the decoding measures (mean effect size =+1.03) in comparison to the nonsignificant effect size for Passage Comprehension of +0.39 (n.s.). Outcomes for the embedded phonics treatment were smaller. Decoding measures averaged an effect size of +0.46 while the effect size for Passage Comprehension was +0.16 (n.s.). Similarly, the regular classroom support treatment had a mean effect size of +0.50 for decoding and +0.17 (n.s.) for Passage Comprehension.

A particularly important additional outcome relates to retentions in kindergarten and first grade. The rate for these very low achieving children was only 9% for *ADD*, 25% for embedded phonics, 30% for regular classroom support, and 41% for control. These differences are statistically significant ($p<.05$). If sustained over time, the savings from not retaining tutored children (i.e., not having to provide an additional year of schooling) could compensate for the considerable expense of the tutoring (see Borman & Hewes, 2003, for a discussion of this issue).

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

This study is particularly important in demonstrating that a multi-year tutoring intervention can greatly increase the decoding performance of struggling readers and greatly reduce retention rates, and that these effects are much larger for a treatment with a strong focus on systematic phonics (specifically *Auditory Discrimination in Depth*) than for an equally intensive tutoring program with less of a phonics emphasis.

Early Steps/Howard Street Tutoring

Early Steps (Morris, Tyner, & Perney, 2000; Santa & Hoien, 1999) is a first grade one-to-one tutoring model that is patterned on *Reading Recovery* but includes more of a focus on phonics than the original *Reading Recovery* program had. Tutors can be the child's own teacher, Title I teacher, or a part-time teacher added to the staff. Tutors receive extensive training, observation, and feedback. They use a series of leveled, non-phonetic books, and follow a schedule of rereading familiar books, word study, sentence writing, and introducing new books, which takes 30 minutes daily during Title I pullout time (so it supplements regular reading instruction).

Morris et al. (2000) evaluated *Early Steps* in mostly high-poverty, majority African American schools in Tennessee. Six schools (n=43) used *Early Steps* with their lowest-achieving first graders. The children were in the lowest 20% of their classes.

Children in 5 matched control schools (n=43) were individually matched on pretests. Control students received small group instruction in their Title I pullout classes; in two cases, the

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

schools used *Direct Instruction*, while the others used unspecified materials. *Early Steps* children scored substantially higher than controls, adjusting for pretests. On an adaption of the Woodcock Word Attack scale, the effect size was +0.92 ($p<.001$), and on an adaptation of the Woodcock Passage Comprehension scale, the difference was +0.80 ($p<.001$), for a mean of +0.86. Effect sizes averaged +1.01 for the lowest-achieving students and +0.86 for the less at-risk children on the two Woodcock measures.

Howard Street Tutoring was a predecessor of *Early* Steps and shares with it all of its essential features. Brown, Morris, & Fields (2005) evaluated an adaptation of the *Howard Street Tutoring* model in seven urban schools in the intermountain west. In three schools that used the tutoring model (n=40), struggling readers were tutored one-to-one twice weekly by teachers or paraprofessionals for 45 minutes over a school year. Teachers generally tutored just one child from their own class. In four matched schools (n=42), struggling readers were given 45 minutes a day of instruction in groups of 2-6, designed to support the *Open Court* basal texts used in all 7 schools. The small groups were also taught by either teachers or paraprofessionals. The students were in grades 2-6, but most were in grades 2-3. 46% of students received free- or reduced-price lunches, 42% were members of minority groups, and 25% were English language learners.

Outcomes strongly favored the children who received one-to-one tutoring. Effect sizes were +0.71 ($p<.01$) on a word recognition test, +0.75 ($p<.01$) on a passage reading test, +0.42 ($p<.05$) on Woodcock Word Attack, and +1.07 ($p<.01$) on Woodcock Passage Comprehension, for an average of +0.74. Comparing children tutored by teachers (n=17) to those tutored by

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

paraprofessionals (n=21), differences favored the teachers (mean ES = +0.47). Averaging across the four measures, the effect size for certified tutors was +1.03 (n=17E, 42C), and this outcome is shown in Table 1. The mean effect size for paraprofessional tutors, +0.55 (n=21E, 42C), is shown in Table 2.

Santa & Høien (1999) evaluated *Early Steps* in four schools in Montana. Two schools (n=23) used *Early Steps* with their lowest-performing first graders, while two matched schools (n=26) provided struggling readers with daily small-group instruction focused on practicing book reading individually and in pairs, with little focus on word study or decoding strategies. At the end of the year, children were individually assessed on experimenter-made measures that did not meet the standards of this review because they may have been inherent to the treatment. However, in fall of second grade, students were given the Woodcock Reading Mastery Test by a school psychologist. Adjusting for pretests, effects on these scales were +0.70 ($p<.05$) for Word Identification, +1.28 ($p<.005$) for Word Attack, and +1.14 ($p<.005$) for Passage Comprehension, for a mean of +1.04.

Reading Rescue

Reading Rescue is a one-to-one tutoring program designed for use either with paraprofessionals or with certified teachers. The emphasis of the tutoring is on phonological awareness, word analysis, comprehension strategies, fluency, and writing, using leveled readers. In a study by Ehri, Dreyer, Flugman, & Gross (2007), language minority students who came

from Spanish-speaking homes in a large city (free lunch=95%) were provided tutoring from December to May of first grade (n=64). Matched students within the same schools or in different schools received either a small group tutoring program, *Voyager Passport* (n=52), or no intervention (n=70). In comparison to the no intervention group, adjusting for pretests, the *Reading Rescue* students scored significantly better ($p<.01$) on Gates MacGinitie Word Decoding (ES=+0.98) and Reading Comprehension (ES=+0.70) for a mean of +0.84. In comparison to the *Voyager Passport* small group intervention, which had the same curricular elements, the effect sizes were +0.79 for Word Decoding and +0.36 for Comprehension, for a mean of +0.57. About half of the tutors were certified teachers and half were paraprofessionals. The adjusted Gates MacGinitie effect size in comparison to untreated controls was +1.08 for certified teachers and reading specialists (n=32), and this is the outcome shown in Table 1. An effect size of +0.89 for paraprofessional tutors appears in Table 2.

Tutoring With Phonology Training

Hatcher, Hulme, & Ellis (1994) compared three variations of one-to-one tutoring in schools in Cumbria, a mostly rural area in the North of England. Six- and seven-year-olds were identified as at risk based on a word reading test; children with low non-verbal IQ scores or who were assigned to special education were excluded. Children were non-randomly assigned to one of four groups, matching on IQ, reading ability, and age. Children in three of the groups received 40 30-minute tutoring sessions over 20 weeks. A total of 23 teachers provided one-to-one

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

tutoring, working with 2-9 children each day in all three variations. In *Reading with Phonology* (n=32), teachers used an adaptation of *Reading Recovery* (as it was at the time) with the addition of instruction focusing on phonological awareness, letter sounds, sound blending, syllables, and words within sentences. Children proceeded through leveled books, did writing activities, and engaged in the sequence of elements typical of *Reading Recovery*. In *Phonological Training Alone* (n=30) children experienced only the phonological training, and did no storybook reading. In *Reading Alone* (n=31), children experienced the leveled book and writing activities adapted from *Reading Recovery* and learned letter names, but the instruction did not include any explicit reference to phonology or letter-sound relationships. The usefulness of context and meaning and the use of self-checking were emphasized. Finally, a control group (n=31) did not receive tutoring.

On immediate posttests, adjusting for pretests, the *Reading + Phonology* children had the highest scores. Effect sizes for *Reading + Phonology* in comparison to the control group were +0.40 ($p<.02$) on British Ability Scales (BAS) Word Reading, +0.62 ($p<.001$) for Neale Accuracy, and +0.94 ($p<.001$) for Neale Comprehension. For *Reading Alone*, effect sizes were +0.05 (n.s.) for BAS Word Reading, +0.38 (n.s.) for Neale Accuracy, and +0.35 (n.s.) for Neale Comprehension. For *Phonology Alone,* the effect sizes were +0.04 (n.s.) for BAS Word Reading, -0.05 (n.s.) for Neale Accuracy, and +0.04 (n.s.) for Neale Comprehension. The Neale measures were given again nine months after the end of the treatment. Effect sizes for *Reading + Phonology* were +0.49 ($p<.01$) for Neale Accuracy and +0.91 ($p<.001$) for Neale

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Comprehension. For *Reading Only* they were +0.21 (n.s.) for Accuracy, +0.31 (n.s.) for Comprehension. For *Phonology Only,* followup effect sizes were -0.18 (n.s.) for Accuracy and +0.07 (n.s.) for Comprehension. The findings strongly support the importance of balancing phonics with reading in context; at both immediate and followup measurement, the combined treatment had by far the most positive effects, both on decoding and on comprehension.

Intensive Reading Remediation

Blachman, Schatschneider, Fletcher, Francis, Clonan, Shaywitz, & Shaywitz (2004) evaluated a one-to-one tutoring program that provided an unusual intensity of services to struggling second and third graders. Children with standard scores below 90 on the Woodcock Basic Skills Cluster or its components (Word Identification and Word Attack) and had IQ's above 80 were identified at the beginning of their Grade 2 or Grade 3 years. Children were then randomly assigned either to tutoring (N=37) or to a control group (N=32). Experimental children received 50 minutes per day of tutoring, 5 days a week, from October to May, receiving an average of 105 hours of instruction in 126 sessions, roughly twice the hours typically received by children tutored in *Reading Recovery*, for example. Tutors were certified teachers, extensively trained in a 45-hour, 15-session training program, and then had 8 additional 2-hour training sessions after tutoring began, as well as regular visits from project leaders to observe tutoring. The sessions followed a regular sequence of review of sound-symbol relationships, manipulating letter cards to practice phoneme analysis and blending, fluency practice, oral reading practice

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

using phonetically controlled text, and dictation. Children proceeded through six levels at their own pace. Control children received whatever remedial reading program was provided by their school, usually a Title I pullout remedial program in groups of from 2 to 8. Control children received an average of 77 hours of small-group remediation in 104 sessions.

After September pretesting, children were posttested in June of the treatment year, and then a followup assessment was administered a year later. Adjusting for pretests, posttest effect sizes were +0.99 for Woodcock Word Identification, +0.96 for Woodcock Word Attack, +0.75 for GORT Oral Reading Quotient, +0.77 for a word reading measure, and +0.80 for a timed measure of word reading efficiency, for a mean of +0.85. On followup measures the effect sizes were +0.81 for Word Identification, +0.46 for Word Attack, +0.47 for GORT Oral Reading, +0.72 for Word Reading, and +0.75 for Word Reading Efficiency, for a mean of +0.64. All differences were statistically significant, $p<.05$ or better.

Targeted Reading Intervention

Targeted Reading Intervention (TRI) is a one-to-one tutoring model designed in particular for isolated rural schools. In it, classroom teachers work individually with struggling readers in kindergarten or first grade for 15 minutes a day. They receive professional development both for tutoring and for reading instruction in general in an initial face-to-face summer institute followed by weekly web conferencing from a university-based consultant to

follow up training, discuss individual children, and, resolve problems. The 1-1 sessions focus on re-reading for fluency (2 min.), word work (6 min.), and guided oral reading (7 min.).

Two randomized quasi-experimental evaluations of TRI were reported by Vernon-Feagans, Amendum, Kainz, Ginsberg, & Bock (2009). In Study 1, 6 elementary schools in two poor rural counties in the southeast were randomly assigned to TRI or control treatments. Students were in kindergarten or first grade, 49% were African American, 33% White, and 10% American Indian, and about 80% qualified for free or reduced-price lunches. Ns were 59E, 66C. On Woodcock Letter-Word Identification, adjusting for pretests, effect sizes were +0.24, and they were +0.25 for Word Attack, for a mean of +0.25.

Study 2 took place in rural Texas and New Mexico. Four schools were assigned at random, in a randomized quasi-experiment, and there were 26 experimental and 17 control classrooms. Children were 37% White, 26% African American, and 35% "other." Children were in kindergarten and first grade (n=97E, 54C). Adjusting for pretests, effect sizes were +0.27 on Woodcock Word Attack, +0.24 on Letter-Word Identification, and +0.50 on Passage Comprehension, for a mean of +0.34.

TEACH

TEACH is a one-to-one tutoring program that focuses on identifying perceptual deficits (such as delayed acquisition of spatial and temporal orientation) using an instrument called SEARCH and then providing one-to-one tutoring focused not on reading instruction but on

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

neurological skills. In a study by Mantzicopoulos, Morrison, Stone, & Setrakian (1992), *TEACH* tutoring was provided to at-risk children in middle class schools in Marin County, California, a San Francisco suburb. Children in the lowest third of their classes on SEARCH at the end of kindergarten were randomly assigned to receive *TEACH* in 50 half-hour, twice-weekly tutoring sessions during first grade (n=59), phonetic tutoring on the same schedule (n=52), or no tutoring as a control group (n=57).

Students were posttested at the end of first and second grades. At the end of first grade, there was a significant effect on Word Attack for the phonetic tutoring group, but not for *TEACH*. Not enough data were given to compute effect sizes. At the end of second grade, posttests adjusted for pretests showed no significant differences between TEACH and control with effect sizes of +0.10 for SDRT Comprehension, +0.09 for SDRT Phonetic Analysis, +0.30 for Woodcock Word Attack, and +0.29 for K-ABC Reading/Decoding, for a mean effect size of +0.19.

An early study of *TEACH* was reported by Arnold, Barnebey, McManus, Smeltzer, Conrad, Winer, & Desgranges (1977). Children in nine first-grade classes in 3 schools (2 inner city, one middle class) were tested with SEARCH in September. 86 children identified as being at risk were matched within schools and assigned non-randomly to *TEACH* (n=23), to regular academic one-to-one tutoring (n=23), or to a no-contact control group (n=40). The two tutored groups received tutoring twice a week for an average of 35 half-hour sessions over 7 months. *TEACH* tutoring focused on specific visual, auditory, and kinesthetic deficits identified by

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

SEARCH, while academic tutoring focused on the reading and math being taught in class. On WRAT-Reading posttests, adjusted for pretests, children in the *TEACH* group gained significantly more than those in the control group (ES=+0.34, $p<.05$), and nonsignificantly more than those in the regular tutoring group (ES=+0.17). On a follow-up test a year later, differences between *TEACH* and control were much larger than at posttest (ES=+0.82, $p<.01$), as were differences between *TEACH* and academic tutoring (ES=+0.77, $p<.01$)

Conclusions: One-to-One Tutoring by Teachers

A total of 19 studies of one-to-one tutoring by teachers met the inclusion criteria. Five studies were randomized and two were randomized quasi-experiments. The overall weighted mean effect size was +0.38. Eight of the studies evaluated *Reading Recovery,* two of which used random assignment, and 13 (4 randomized and 2 randomized quasi-experiments) evaluated other teacher tutoring models. The outcomes for *Reading Recovery* were positive, but less so than might have been expected. Across all 8 studies, the weighted mean effect size was only +0.23. This mean did not include measures from the Clay Diagnostic Survey (such as Text Reading Level), which were considered inherent to the treatment and were therefore excluded. This exclusion of the Diagnostic Survey measures and the inclusion of all students who received *Reading Recovery* tutoring (not just successful "discontinued" students) explains the difference between effect sizes reported here and those of other reviews.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

The two randomized studies of *Reading Recovery* found conflicting outcomes. The Ohio statewide study (Pinnell et al., 1994) reported effect sizes on Gates MacGinitie of +0.19 for schools in which *Reading Recovery* teachers were experienced with the program, in comparison to students who did not receive tutoring or any other intervention. In contrast, an earlier randomized study by Pinnell, DeFord, & Lyons (1988) found an effect size of +0.52 on CTBS.

Reading Recovery effects were particularly weak in evaluations in large urban districts in the U.S. In large studies in Austin, Texas (Williams et al., 1995) and Detroit (Huggins, 1999), control groups scored slightly better than did *Reading Recovery*-tutored students. In a study by Acalin (1995), students in *Project Read,* a phonetic small-group intervention, scored slightly better than did students tutored one-to-one in *Reading Recovery*. In contrast, all three of the non-U.S. studies found strong positive effects. Burroughs-Lange (2008) found an effect size of +0.76 in London schools, Hurry & Sylva (2007) found an effect size of +0.85 in English primary schools, and Center et al. (1995) reported an effect size of +0.86 in urban Australian schools. The difference between the American studies and the UK and Australian studies may relate to the nature of the control groups, which were typically receiving small group, phonetic instruction in the U.S. but were less likely before the late 1990's to receive special treatment in the UK or Australia.

The findings for *Reading Recovery* were particularly surprising in light of the much larger effects for other one-to-one tutoring programs, which had a weighted mean effect size of +0.46 across 13 studies. One possible explanation relates to the role of phonics. The 9 studies of

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

tutoring programs with a strong emphasis on phonics had a mean effect size of +0.69 (in contrast to +0.23 for less phonetic approaches, *Reading Recovery* and *TEACH*). Within-study comparisons also point to the importance of a phonics emphasis; Hatcher et al. (1994) and Iverson & Tunmer (1993) explicitly added phonics instruction to *Reading Recovery*-like tutoring models, and in both cases effect sizes were higher in the combined treatments. Since *Reading Recovery* now has more of a phonics focus than it did when most evaluations were done, it is unclear how these findings relate to the current program. The role of phonics in tutoring outcomes is discussed further later in this article.

Tutoring by Paraprofessionals and Volunteers

One-to-one tutoring by certified teachers is expensive, and in high-poverty communities with shortages of teachers, allocating scarce qualified teachers to small numbers of children may be hard to justify. For those reasons, many schools have long used paraprofessionals or volunteers as tutors, usually with materials specifically designed for this purpose. Ritter et al. (2009) and Wasik (1997) reviewed research on volunteer tutoring programs, and both reported substantial positive effects. Table 2 summarizes research on the reading outcomes of one-to-one tutoring programs in which the tutors are paraprofessionals or volunteers.

==========

TABLE 2 HERE

==========

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Sound Partners

Sound Partners (Vadasy, Wayne, O'Connor, Jenkins, Pool, Firebaugh, & Peyton, 2005) is a one-to-one tutoring program that provides struggling first graders with 30 minutes of daily tutoring by non-teacher tutors. It consists of 100 structured lessons focusing on phonological awareness, phonics, word identification, text reading, and writing.

Jenkins, Peyton, Sanders, & Vadasy (2004) compared two adaptations of *Sound Partners* to a control group in a study of at-risk first graders. Children scoring below the 25th percentile on the WRAT-R (1984) were randomly assigned within 11 urban schools in the Northwest to two tutoring variations, one using more decodable texts in which 85% of words children encountered in storybooks were decodable using letter sounds they had been taught (n=39), and one using less decodable texts in which only 11% of storybook words were decodable (n=40). In both conditions children were tutored by paraprofessionals 4 days a week for 25 weeks. Tutors received 3 hours of training and were then visited by research staff weekly for coaching and followup. The lessons were scripted and focused on letter-sound correspondence, sound blending, and storybook reading. For comparison, 20 matched children were non-randomly assigned to a control condition, in which no tutoring was provided.

On an extensive battery of assessments, there were no differences between the groups tutored with more decodable or less decodable texts. Data from these variations are therefore combined and compared to data from the control (untutored) condition. Adjusting for pretests, posttest effect sizes were +0.50 for Woodcock Word Identification, +0.77 for Woodcock Word

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Attack, +1.13 for the Bryant, +0.38 for TOWRE Phonemic Decoding, +0.74 for WRAT-Reading, and +0.52 for TOWRE Sight Word Reading, for a mean of +0.67 for the six combined decoding and word reading measures. The effect size for Woodcock Passage Comprehension was +0.81, for an overall mean across all 7 measures of +0.69. All differences were statistically significant ($p<.05$ or better) except TOWRE Phonemic Decoding.

Mooney (2003) evaluated *Sound Partners* in seven elementary schools in a Midwestern city with first graders identified as at risk for emotional and behavioral disorders (EBD). The children were 66% boys, and 55% qualified for free or reduced-price lunches. 68% were White, 21% African-American, and 9% Hispanic. They were randomly assigned to *Sound Partners* (n=28) or control (n=19) within schools. All students used the phonics component of *Open Court Reading* in their regular reading classes, and control students received a home-school intervention called *First Step to Success* intended to improve their social skills.

Children were individually pre- and posttested on Woodcock and DIBELS. Despite random assignment, pretest differences favored the experimental students (pre ES=+0.48 for Woodcock, +0.50 for DIBELS), barely meeting inclusion standards. Adjusting for these differences, Woodcock effect sizes were +0.26 for Basic Skills, +0.22 for Comprehension, and +0.21 for Total Reading. For DIBELS, effect sizes were +0.54 for Nonsense Word Fluency and +0.35 for Oral Reading Fluency. Averaging Total Reading and DIBELS, the effect size was +0.34.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

An adaptation of *Sound Partners* was evaluated with struggling readers in grades 2-3 by Vadasy, Sanders, & Tudor (2007). Paraprofessionals provided children scripted lessons 30 minutes per day, 4 days a week, for 15 weeks. A control group received no tutoring. The children were nominated by 26 teachers in nine urban schools in the Northwest, and scored between the 10th and 37th percentiles at pretest. A total of 43 children (23E, 20C) were randomly assigned to treatments. On a reading accuracy composite score derived from Woodcock Word Attack and Word Identification scales, students who received tutoring scored significantly higher than controls, adjusting for pretests (ES=+0.49, p<.008). On an adaptation of DIBELS-Fluency, the effect size was +0.55 (p<.014). The mean effect size was +0.52.

Vadasy, Sanders, & Peyton (2005) evaluated *Sound Partners* with first graders scoring in the lowest quartile of their classes. A total of 57 children (19 in each group) in a Northwest city were matched on pretests in two variations of *Sound Partners* or a control group that did not receive tutoring. One variation, *Reading Practice*, emphasized oral reading of text in addition to *Sound Partners* phonetic tutoring. The other, *Word Study*, added additional tutoring on decoding words. All tutored children received 30-minute sessions 4 times a week from October to May. Adjusting for pretests, effect sizes comparing the two *Sound Partners* groups to controls averaged +1.02 on WRAT Reading, +1.06 on Woodcock Word Attack, +0.86 on Woodcock Word Identification, +0.66 on Woodcock Passage Comprehension, +0.55 on TOWRE Phonemic Decoding, +0.56 on TOWRE Sight Word, +0.33 on Passage Reading Rate, and +0.66 for Passage Reading Accuracy, for a mean of +0.71. There were no differences between the two

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

variations except on Passage Reading Rate (ES=+0.52) and Accuracy (ES=+0.36) in favor of the *Reading Practice* treatment.

Across 4 studies, the mean effect size for *Sound Partners* was +0.56.

The Reading Connection

Compton (1992) carried out a randomized evaluation of a program in which juniors and seniors at Western Michigan University provided one-to-one tutoring to low-achieving first graders in Kalamazoo. Children eligible for Title I were randomly assigned to tutoring (n=266) or to small-group Title I services (n=217). The tutoring method was based on *Reading Recovery*. 53% of the children were members of minority groups. No pretests were given but random assignment using a computerized random number generator made initial equality likely. At posttest in spring of first grade the effect size on the Iowa Test of Basic Skills was +0.22 ($p<.05$).

SMART

Start Making a Reader Today (SMART) is a volunteer tutoring program designed to provide training and supervision to help community volunteers succeed in tutoring struggling readers in grades 1-2. Children nominated as at-risk by their teachers are tutored 30 minutes a day twice a week through first and second grade. Tutors are given a handbook with simple strategies for reading with children. These focus on letter-sound relationships, predictions, reading stories in many genres, and asking questions about core elements of the texts.

Baker, Gersten, & Keating (2000) evaluated *SMART* in a randomized experiment. First graders in 24 classrooms in 6 schools were randomly assigned to *SMART* (n=43) or a no-tutoring control (n=41). The children were in their respective treatments for two years. At the end of first grade, adjusting for pretests, children in the *SMART* group scored much higher than controls on Woodcock Word Identification (ES=+0.60), Woodcock Passage Comprehension (ES=+0.47), and Oral Reading Fluency (ES=+0.52), for a mean of +0.54. At the end of second grade, differences again favored the *SMART* students on Woodcock Word Identification (ES=+0.62), Passage Comprehension (ES=+0.36, p<.067), and Word Comprehension (ES=+0.46, p<.025), as well as Oral Reading Fluency (ES=+0.54, p<.014). The average of the second-grade effect sizes is +0.50.

Edmark Reading Program

The *Edmark Reading Program* is a highly-structured one-to-one tutoring program designed to build a 150-word vocabulary in beginning or disabled readers. Mayfield (2000) evaluated *Edmark* with low-achieving (but not special education) first graders in three schools in rural northern Louisiana. For the first semester, children were randomly assigned within schools to either receive *Edmark* tutoring from America Reads volunteers (n=31) or to be in a control group that was read to daily in small groups (n=29). On the Woodcock Passage Comprehension test, adjusted for pretests, the *Edmark* students scored significantly higher than controls

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

(ES=+0.51, p<.017), but there were no effects on Letter-Word Identification (ES=+0.14, n.s.) or on Word Attack (ES=+0.03, n.s.), for a mean of ES=+0.23.

Wallach and Wallach

The *Wallach and Wallach* tutoring program was an early phonetic approach for struggling first graders. Paraprofessionals use the program a half hour each day throughout first grade. A study in rural Roanoke Rapids, North Carolina (Dorval, Wallach, & Wallach, 1978) compared three groups of first graders, all performing below the 36th percentile on CTBS pre-reading at the end of kindergarten. Children who received *Wallach and Wallach* tutoring (n=20) were compared to matched children in the same school (n=20), and matched children in a different school (n=18). Effect sizes on CTBS Total Reading at the end of first grade, adjusting for pretest differences, were +0.66 compared to same-school controls, and +0.77 compared to different-school controls, for a mean of +0.71.

Programmed Tutorial Reading

Programmed Tutorial Reading (PTR) was a Houghton-Mifflin tutoring program for struggling readers. Paraprofessional tutors were given step-by-step procedures for a series of lessons that children proceeded through at their own levels and rates. The curriculum focus was on word attack and comprehension skills. McGrady (1984) evaluated *PTR* in a rural/suburban district near Lafayette, Indiana. Fourth graders scoring below the 37th percentile were given 15

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

minutes of tutoring daily (n=35). Matched control students (n=34) within the same six schools did not receive remediation. On ITBS-Comprehension tests, tutored students gained non-significantly more than controls, adjusting for pretests (ES=+0.20).

Reading Rescue: Paraprofessional Tutors

As described previously, Ehri et al. (2007) compared effects of *Reading Rescue* as taught by certified teachers and reading specialists (ES=+1.08; n=32E, 70C) to those achieved by paraprofessionals (ES=+0.89; n=26E, 70C). See the description under One-to-One Tutoring by Teachers, above.

Howard Street Tutoring: Paraprofessional Tutors

The *Howard Street Tutoring* model was evaluated by Brown et al. (2005) with both certified and paraprofessional tutors. Effect sizes for certified teachers on four measures averaged +1.03 (n=17E, 42C) and for paraprofessionals they averaged +0.55 (n=21E, 42C).

Tutoring by Volunteers

Experience Corps

Experience Corps is a large national program that brings adults over age 55 into elementary schools to tutor at risk children. The tutors are trained to use a structured curriculum, whose details vary from place to place, and they receive training and monitoring from program

coordinators. Tutoring sessions are provided approximately 2-3 times a week over the course of the school year.

Morrow-Howell, Jonsen-Reid, McCrary, Lee, & Spitznagel (2009) carried out a large, randomized evaluation of *Experience Corps* in a total of 23 schools in Boston, New York City, and Port Arthur, Texas. Children in grades 1-3 who had been identified by their teachers as struggling in reading were randomly assigned to tutoring or no-tutoring conditions within schools. There were 881 students (430E, 451C). The students were very disadvantaged and low achieving; 94% received free lunch, 58% were African American, and 36% were Hispanic, with 24% limited English proficient. On Woodcock Word Attack, adjusted for the pretest score and covariates including gender, ethnicity, grade, program site, classroom behavior, IEP, and LEP, the effect size was +0.10 (p=.07). The Woodcock Passage Comprehension effect size was +0.13 (p=.04), for a mean of ES=+0.11. Effect sizes were somewhat higher for students who recived at least 35 tutoring sessions (mean ES=+0.15).

Book Buddies

Book Buddies is a volunteer tutoring program in which struggling first graders receive 45-minute tutoring sessions 2-3 times a week for about 10 weeks, a total of 40 lessons. The program includes re-reading of familiar books, word study, writing, and introducing new books. Meier & Invernizzi (2001) evaluated *Book Buddies* in a school in the South Bronx, New York City. The tutors were retired volunteers from the community. Tutees were first graders who

scored in the lowest 25% of their grades at the end of kindergarten. 99% of children received free lunches, 69% were Hispanic, and 30% were African American. Children were randomly assigned to tutoring (n=28) or control (n=27) for a 4-month experiment, with pretesting in September and posttesting in January. On WRAT-Reading, adjusting for pretests, the posttest effect size was +1.00 ($p<.05$), and on the number of words read correctly in one minute, differences were +0.78 ($p<.05$), for a mean of +0.89.

HOSTS

HOSTS, for *Help One Student to Succeed*, is a volunteer tutoring program that uses diagnostic testing to direct tutors to appropriate materials for each child. A study by Ramey (1991) in Seattle schools compared a variety of approaches for low-achieving students in grades 2-5. Students in *HOSTS* (n=18) were compared to matched students in a traditional pullout program (n=220). At the end of the treatment year, there were no differences on CAT Reading (ES=+0.05, n.s.), and in a followup year there were again no differences (ES= -0.01, n.s.).

Other Volunteer Tutoring

A study by Ritter (2000) provides a cautionary note about the effects of volunteer tutoring. The West Philadelphia Tutoring Project (WPTP) was a program in which University of Pennsylvania students worked one on one with students in grades 2-5 once a week for a year. Most tutors were given little guidance and no specific materials to use with their tutees, and

across 11 participating schools practices ranged widely. Often, tutors worked with students on their homework or made up activities to try to help their tutees in reading and math. Only in a few schools did the tutoring coordinator provide curriculum guides for the tutors to use. In a year-long evaluation, 385 students (196E, 189C) were randomly assigned within schools to tutored or control conditions. 87% of the students received free lunches, and 96% were African American. The children were identified by their teachers as being low achievers. On SAT-9 posttests, adjusted for pretests and other variables, the effect size was -0.10 (n.s.). Outcomes were similar for math test scores and for reading and math grades.

Pullen et al. (2004) evaluated a volunteer tutoring model for struggling first graders that used university students as tutors. Tutors were trained and then regularly observed using a model in which children read familiar books, were assessed using a running record, and then were introduced to a new book. The emphasis was on repeated reading and coaching rather than decoding or phonics. In a 3-month experiment, first graders who scored below the 30th percentile on an invented spelling test in 10 schools in north-central Florida were randomly assigned to tutoring (n=23) or non-tutoring (n=24) conditions. The tutors, mostly education majors, were given 4 hours of training as well as lesson guides and student materials, and were then observed frequently by supervisors. Tutored students scored significantly better than controls on Letter Identification (ES=+0.23, $p<.05$) and Word Attack (ES=+0.80, $p<.05$).

Rimm-Kaufman, Kagan, & Byers (1999) evaluated a volunteer tutoring program in which adult volunteers, mostly retired people (half of whom were retired teachers) in Cambridge,

Massachusetts, tutored struggling first graders. Tutored children were assigned to receive tutoring 3 times a week for 45 minutes from October to May. The tutoring model emphasized phonics in context, reading comprehension, and reading for meaning, using shared book reading, games, writing, and other activities. Children were randomly assigned to tutoring (n=21) or untreated control (n=21) conditions. Overall, children were 29% African American, 26% Haitian Creole, and 26% White, and 60% received free or reduced-price lunches. Clay's Observation Survey was used as a pre- and posttest. Tutored children scored marginally higher than controls on reading level (ES=+0.35, p=.08) and nonsignificantly higher on a word knowledge test (ES=+0.18), for a mean of +0.27.

Allor & McCathren (2004) evaluated a tutoring program for first graders that provided volunteer college student tutors with minimal training. The students were America Reads members or unpaid volunteers who were all education majors at a university in a Southern city. The tutors received three one-hour training sessions, as well as assistance on site. First grade tutees received tutoring for 15-20 minutes 4 times per week, usually with two different tutors. Tutors had detailed guidelines to follow that led children through games and activities focusing on phonemic awareness, decoding, and reading of leveled books, and comprehension.

Two successive cohorts were studied. In the first, eight elementary schools were involved. All students received free or reduced-price lunches, and almost all were African-American. The four lowest-achieving students in each class were identified and randomly assigned to tutoring (N=61) or non-tutoring (N=25) conditions. Adjusting for pretests, end of

year posttest effect sizes averaged +0.59 on Woodcock Word Identification, +0.93 for Woodcock Word Attack, +0.49 for Woodcock Passage Comprehension, +0.41 for TOWRE Real Words, and +1.44 for TOWRE Non Words, for a mean of +0.77. All differences except TOWRE Real Words were significant ($p<.05$). In a second cohort, procedures were similar, except that 6-8 low-achieving students were chosen from each class. There were 10 schools (N=76E, 81C) with 94% of students receiving free or reduced-price lunches and 96% African-American. Effect sizes were smaller than in cohort 1:+0.11 (n.s.) for Word Identification, +0.80 ($p<.001$) for Word Attack, -0.16 (n.s.) for Passage Comprehension, +0.14 (n.s.) for TOWRE Real Words, +0.61 ($p<.01$) for TOWRE Non Words, and +0.31 ($p<.05$) for DIBELS Nonsense Word Fluency, for a mean effect size of +0.30. Averaging across the two cohorts yields a mean effect size of +0.54.

Conclusions: One-to-One Tutoring by Paraprofessionals and Volunteers

A total of 11 studies (6 randomized) evaluated one-to-one tutoring by paraprofessionals, with a weighted mean effect size of +0.38. Effect sizes across 7 studies of volunteer tutoring (4 fully randomized) averaged +0.16. However, this mean was brought down by the large sample sizes and low effect sizes of the studies of *Experience Corps* and the *West Philadelphia Tutoring Program*, which provided tutoring only 1-2 times a week. Excluding these, the weighted mean effect size for 5 studies of volunteer tutoring was +0.51. These five small studies had better-qualified and better-trained paraprofessionals and volunteers than might be common.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

The effects seen for paraprofessional tutors and for volunteer tutors using structured and intensive programs pose a real challenge to the idea that only certified teachers can be effective tutors. The overall effect size for the 18 studies, ES=+0.24, is higher than the mean for *Reading Recovery,* although much lower than the effects for phonetic tutoring by teachers (ES=+0.50). It is important to note that Brown et al. (2005) and Ehri et al. (2007) compared certified and paraprofessional tutors and found the certified tutors to be more effective, but both studies still found substantial positive effects for the paraprofessional tutors. What these findings imply is that schools might use a mix of certified, paraprofessional, and volunteer tutors, using the certified teachers as leaders and to work with the most difficult children. One-to-one tutoring is clearly very effective, and when resources are limited, well-structured programs making use of paraprofessionals and volunteers may reach more struggling readers for the same cost as serving many fewer children with certified teachers. The comparison between certified and paraprofessional and volunteer tutors is discussed further later in this article.

Small Group Tutorials

The most common form of remedial or supplementary instruction for struggling readers is additional teaching in small groups, typically 30-45 minutes daily. Since the 1960's, small group remediation has been the predominant use of Title I funds and it is the most common format for children with reading disabilities. Small group tutorials are potentially more cost-effective than one-to-one tutoring from teachers, because several children are taught at the same

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

time, and the group setting creates possibilities for children to learn from each other as well as from the teacher. On the other hand, small group teaching can be more of the same type of instruction that did not work the first time, can be difficult to coordinate with regular classroom instruction, and does not allow teachers to tailor instruction to students' needs as much as one-to-one instruction does.

Table 3 summarizes research on the reading outcomes of participation in small-group tutorials.

==========

TABLE 3 HERE

==========

Corrective Reading

Torgesen et al. (2006) evaluated a variation of *Direct Instruction* for struggling readers called *Corrective Reading* in a large randomized evaluation of four programs: *Corrective Reading, Failure-Free Reading, Wilson Reading,* and *Spell Read*. A total of 32 schools around Pittsburgh, PA, were randomly assigned to participate in one of the four programs, and then students scoring above the 5th percentile on PPVT and below the 30th percentile on TOWRE were randomly assigned within schools to experimental or control groups. Experimental and control students were well matched on pretests and demographics. They averaged 80% White and 20% African American, and 44% received free lunch. Averaging across 7 Woodcock,

TOWRE, AIMSweb and GRADE scales, the average effect size for *Corrective Reading* was +0.22 for third grade (n=79) and +0.09 for fifth grade (n=86), for an average of +0.16. A one-year followup found that this difference had diminished to an effect size of +0.06. Results for *Spell Read* and *Wilson Reading* appear later in this section.

In a study in Australia, Hempenstall (2008) evaluated the use of *Corrective Reading: Decoding* with mid- to upper-primary students, mostly ages 8 to 11, who were having serious difficulties in learning to read. The students were in 5 state and 4 Catholic schools in a suburb of Melbourne with a high percentage of disadvantaged students. In a 7-month matched study, students were either taught *Corrective Reading* in groups of 10 (n=134), or remained in a waitlist control group that did not receive a remedial program (n=72). The *Corrective Instruction* sessions were provided in 50-minute pullout sessions 5 days per week. Students received 60-65 lessons over 5-10 months. On Woodcock Word Attack measures, adjusted for pretests, the effect size was +1.22 ($p<.001$).

Spell Read

Kaplan Spell Read is a program for struggling readers that emphasizes systematic and explicit fluency-oriented instruction in phonics. Students engage in shared reading and discussion of leveled books, as well as writing and spelling activities.

Spell Read was one of the four programs evaluated by Torgesen et al. (2006), along with *Failure-Free Reading, Wilson Reading*, and *Corrective Reading*. A total of 32 schools were

randomly assigned to one of these programs and then students who scored above the 5th percentile on PPVT and below the 30th percentile on TOWRE were randomly assigned within each school to experimental or control conditions. Overall, approximately 69% of students were White and 31% were African American, and 44% qualified for free lunch. Averaging across 7 Woodcock, TOWRE, AIMSweb, and GRADE measures, third graders (n=92) in *Spell Read* classes scored better than controls (ES=+0.21). Differences were smaller for fifth graders (n=104) (ES=+0.12), and the average was +0.17. A one-year follow-up found an average effect size of +0.11 (Torgesen et al, 2007).

Wilson Reading

The Wilson Reading System is a complete curriculum for teaching reading to children in grades 3 and above who did not learn to read well in the primary grades. It focuses on phonics, but also emphasizes fluency, vocabulary, and comprehension. It uses a multisensory approach based on Orton-Gillingham methods. Teachers receive extensive professional development, including a ten-hour in-service at the beginning and participation in an on-line academy for ongoing assistance and materials.

A randomized evaluation of *Wilson Reading* was carried out by Torgesen et al. (2007) in schools around Pittsburgh. A total of 32 school units were randomly assigned to participate in one of four programs, and students in grades 3 and 5 were randomly assigned to treatment or control conditions. To be eligible, the students had to score above the 5th percentile on the PPVT

and below the 30th percentile on the TOWRE. Experimental and control students were fairly well matched on a battery of baseline assessments as well as demographic factors, and averaged 56% White and 44% African American. Controls were higher in percent free lunch (56% vs 37% in third grade). There were 51 *Wilson Reading* and 19 controls in third grade, and 52 *Wilson Reading* and 36 controls in fifth grade.

After adjustments for pretests, students who experienced *Wilson Reading* generally scored higher than controls in both grades. Averaging across 7 measures, the mean effect size was +0.26 for third grade and +0.08 for fifth grade, for a mean of +0.17. At both grade levels, significant ($p < .05$) positive effects were found on Word Attack, but there were no significant differences on GRADE or Woodcock Passage Comprehension measures at either grade level. A one-year follow-up found an average effect size of +0.17.

Failure Free Reading

Failure Free Reading is an intervention program for low-achieving readers that uses a balance of structured teacher instruction, computers, and video to build students' phonics comprehension, fluency, and vocabulary skills. Children are frequently assessed and given instruction at their own levels.

Torgesen et. al. (2007) evaluated *Failure Free Reading* as part of a randomized study of four programs in schools near Pittsburgh. A total of 32 schools were randomly assigned to each treatment, and then third and fifth graders were randomly assigned to treatment and control

groups within each school. Students had to score above the fifth percentile on PPVT and below the 30th percentile on TOWRE, so the students were moderately low in achievement. *Failure Free Reading* and control students were fairly well matched on demographics and pretests and were about 20% African American and 80% white, with about 44% qualifying for free lunch.

After adjustments for pretests, outcomes varied for the third and fifth graders. Among third graders, *Failure Free Reading* students (n=51) scored better than controls (n=38) (ES=+0.10 across 7 measures). Differences on 4 of the measures were significant ($p < .05$). However, fifth graders in *Failure Free Reading* (N=62) and control (N=66) did not differ (ES = -0.00, n.s.). The mean effect size across grades was +0.05. A one-year follow-up found an average effect size of +0.14.

QuickReads

QuickReads is a supplemental program for struggling readers that provides tutoring to dyads that focuses on repeated reading, letters and sounds, and comprehension. In a study by Vadasy & Sanders (2008), second and third graders (n=82E, 80C) were randomly assigned in pairs to be tutored 30 minutes a day, 4 days a week, for 15 weeks, mostly by paraprofessionals, or to remain in a control group that did not receive tutoring. Students were 30% White, 28% African American, 23% Hispanic, and 16% Asian. Effect sizes, adjusted for pretests, were +0.27 for Woodcock Word Identification, +0.12 for TOWRE Sight Word Efficiency, +0.16 for GORT Comprehension, +0.30 for GORT Fluency, and +0.27 for DIBELS, for a mean of +0.22.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

An evaluation of *QuickReads* with fourth and fifth grade students who were reading below grade level was reported by Vadasy & Sanders (2008). The students were from 12 schools in a large city in the Northwest. 10% of students were African American, 25% White, 12% Hispanic, and 9% Asian. Students were randomly assigned to dyads and then dyads were randomly assigned to *QuickReads* (n=54) or control (n=65). The *QuickReads* students were tutored in pairs by paraprofessionals for 30 minutes a day, four days a week, for 20 weeks. Control students received no supplemental instruction. Adjusting for pretests, effect sizes were +0.33 for Woodcock Word Identification, +0.21 for TOWRE Sight Word, +0.09 for Woodcock Word Comprehension, +0.23 for Woodcock Passage Comprehension, and +0.16 for DIBELS Fluency, for a mean of +0.20.

Targeted Intervention

Wang & Algozzine (2008) evaluated a small group targeted intervention program for struggling first graders. The program was based on the formats and sequence of skills from *Direct Instruction,* and provided 110 daily 10-15 minute lessons to supplement a 120-minute *Open Court* reading period, used in both experimental and control schools. The supplemental lessons were taught by teaching assistants. Students were in 6 urban schools in which 80% of students received free lunch and 89% were African American or Hispanic. Schools were randomly assigned to conditions, with 4 schools (n=101) assigned to the targeted intervention and 2 (n=38) to control, which continued with the core curriculum. Adjusting for pretests, effect

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

sizes were +0.32 (p<.03) for Woodcock Word Identification, +0.43 (p<.02) for Woodcock Word Attack, +0.13 (n.s.) for Woodcock Passage Comprehension, and -0.12 (n.s.) for DIBELS Nonsense Word Fluency, for a mean of +0.19.

Proactive and Responsive Reading

Mathes, Denton, Fletcher, Anthony, Francis, & Schatschneider (2005) carried out an evaluation of two conceptually distinct approaches to small-group supplemental instruction for struggling first graders. Both approaches used 3-member tutoring groups. One, called *Proactive Reading*, is based on *Direct Instruction* and emphasizes phonemic awareness, letter sounds, reading of decodable text, fluency, and comprehension of connected text. The teaching emphasizes rapid instruction, frequent opportunities to respond, positive feedback, and immediate error correction. The second, called *Responsive Reading*, placed less of a focus on phonemic awareness and decoding, and had less of a structured sequence of steps in each lesson. Teachers alternated among children to provide intensive scaffolding at each child's level. A daily lesson cycle consisted of fluency building, assessment, letters and words, supported reading, and supported writing. Teachers chose leveled books that did not have decodable text.

The small-group interventions were implemented for 40 minutes daily over the course of first grade. Low achieving students in six schools in a large urban district in Texas were randomly assigned to *Proactive Reading* (n=80), *Responsive Reading* (n=83), or classroom instruction only (n=82). On Woodcock Johnson III posttests, effect sizes for *Proactive Reading*,

compared to controls, were +0.76 (p<.001) for Word Attack, +0.51 (p<.001) for Word Identification, and +0.21 (n.s.) for Passage Comprehension, for a mean of +0.49. For *Responsive Reading*, effect sizes in comparison to controls were +0.28 (n.s.) for Word Attack, +0.36 (p<.05) for Word Identification, and +0.30 (p<.10) for Passage Comprehension, for a mean of +0.31. Differences between *Proactive Reading* and *Responsive Reading* were significant for Word Attack (p<.05) but not for Word Identification or Passage Comprehension.

New Heights Reading Program

The *New Heights Reading Program* is a small-group remedial program for students reading at least nine months below their grade level. Originally developed in New Zealand under the name *Rainbow Reading, New Heights* was adapted for the U.S. Students advance at their own paces through a series of leveled readers and use audiotapes and text-based activity sheets. Teachers conference individually with students to review new vocabulary, encourage progress, and build motivation.

Lesnick (2006) evaluated *New Heights* in an 18-week study in two very different school districts, Philadelphia and Lower Merion, Pennsylvania. Students in the Philadelphia schools were almost all African American or Hispanic, and all received free lunches. In Lower Merion, a Philadelphia suburb, students were 72% White, and only 6% received free lunches. Third and fifth graders in nine schools who were reading at least 9 months below grade level were randomly assigned within schools to New Heights (n=30 classes, 118 students) or to a control

group (n=29 classes, 115 students), in which most students also received some sort of remediation, such as computer programs, tutors, or small group instruction. On posstests adjusted for pretests and demographic variables, there were no significant differences on DIBELS (ES= -0.01), TOWRE (ES= -0.02), or DRP (ES=+0.09), for a mean of +0.02. Results were nearly identical in Philadelphia and Lower Merion, in third and fifth grades.

Read Naturally

Read Naturally is a small group supplementary program that focuses on building fluency among low achievers. Students start with a "cold read" of a story and then practice with audiotapes until they reach a fluency target. They then answer comprehension questions and retell the story in writing.

An evaluation of *Read Naturally* was carried out by the Minneapolis Public Schools (Heistad, 2005). This evaluation combined students taught *Read Naturally* during the school day (as the program is intended to be used), and those using it as an after school program. The developer provided data on a school that used *Read Naturally* during the school day, and those data are used here. Each *Read Naturally* student was individually matched with a control student in third and fifth grade. Control students did not receive any supplemental instruction. On posttests, adjusted for pretests, effect sizes were +0.21 for the NALT and +0.34 for the Minnesota Comprehensive Assessments, for a mean of +0.27.

Voyager Passport

Voyager Passport is a commercial small-group program for struggling readers that emphasizes phonics, phonemic awareness, comprehension, vocabulary, and fluency in daily 30-40 minute sessions. The lessons were scripted for teachers. Parental follow-up was encouraged and take-home readers were provided. In a study of the *Reading Rescue* one-to-one tutoring model by Ehri et al. (2007), first graders using *Voyager Passport* (n=52) were compared to matched controls receiving no intervention (n=70). Tutees were Spanish-dominant language minority students; 95% received free lunches. Adjusting for pretests, the *Voyager Passport* students scored non-significantly better than those receiving no intervention on Gates Word Decoding (ES=+0.26) and Reading Comprehension (ES=+0.36), for a mean of +0.31. However, they scored much lower than students receiving the *Reading Rescue* one-to-one tutoring program (mean effect size
= -0.57)

Empower Reading

Lovett, Lacerenza, Borden, Frijters, Steinbach, & DePalma (2000) carried out a study that compared alternative approaches to small-group remediation in research classrooms in Toronto. Children ages 6 to 13 who were reading at least 2 standard deviations below age norms were identified. They were randomly assigned to one of four treatments, each of which involved 70 hours of instruction in groups of 3. One was a phonetic approach using materials from *Direct*

Instruction (n=18). One was a metacognitive approach, in which children were taught to use metacognitive strategies to decode and comprehend text (n=20). A third was a combination of the two (n=15), which was later given the name *Empower Reading.* A control treatment focused on math and self-help skills (n=22). Adjusting for pretests, the combined approach had effect sizes of +1.23 for Woodcock Word Attack, +0.45 for Woodcock Word Identification, +0.49 for Woodcock Passage Comprehension, and +0.68 for WRAT – Reading, for a mean of +0.71. The corresponding effect sizes for the phonics-only approach were +0.62 for Word Attack, +0.37 for Word ID, +0.52 for Passage Comprehension, and +0.29 for WRAT, for a mean of +0.45. For metacognitive training only, effect sizes were +0.60 for Word Attack, +0.41 for Word ID, +0.61 for Passage Comprehension, and +0.57 for WRAT, for a mean of +0.55.

A second study, by Lovett, DePalma, Frijters, Steinbach, Temple, Benson, & Lacerenza (2008), evaluated the same three treatments, but analyzed them together as variations of one overall intervention. Groups of 4-8 children who were performing at least one standard deviation below grade level were randomly assigned to treatment or control conditions in 16 Toronto schools. Children were in grades 2-8, with an average age of 10.2 at program entry. 46% of the students spoke a language other than English at home, most frequently Portuguese or Spanish. Students assigned to the experimental group (n=122) received tutoring in small groups in their schools, participating in an average of 105 hours of instruction. Control students (n=44) participated in their schools' special education program. Posttests adjusted for pretests showed effect sizes of +0.40 ($p<.01$) for WRAT-Reading, +0.22 (n.s.) for Woodcock Word

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Identification, +0.37 (p<.01) for Woodcock Word Attack, and +0.02 (n.s.) for Woodcock Passage Comprehension, for a mean of +0.25.

Schools and Homes in Partnership (SHIP)

Schools and Homes in Partnership (SHIP) was a program that provided 30 minutes of daily supplemental instruction to struggling readers in groups of 2-3, over a two-year period. It used the *Direct Instruction Reading Mastery* program in grades 1-2 and *Corrective Reading* in grades 3-4, focusing on explicit instruction in phonics and fluency. Instructional assistants from a mostly Hispanic community were given 10 hours of training to provide the supplemental instruction and were then observed and given feedback on a regular basis. A randomized evaluation of *SHIP* was reported by Gunn, Smolkowski, Biglan, & Black (2000). In it, low-achieving children in grades K-3 in three rural districts in central Oregon were randomly assigned within schools to *SHIP* (N=105) or control (N=106) conditions, stratifying on age, ethnicity, and achievement level. 62% of children were Hispanic, and 38% were White. Students were pretested in the fall of the year they were identified and then posttested each spring for two years. At the end of two years, controlling for pretests, effect sizes were +0.31 (p <.02) for Woodcock Letter-Word Identification, +0.66 (p<.001) for Woodcock Word Attack, +0.20 (p<.02) for Woodcock Vocabulary, +0.29 (p<.02) for Woodcock Comprehension, and +0.24 (p<.06) for DIBELS Oral Reading Fluency, for a mean of +0.34. Follow-up testing two

years later found an effect size of +0.37 for Hispanics and +0.08 for non-Hispanics, for a weighted mean of +0.25.

Small-Group Phonics Tutoring

Gottshall (2007) developed and evaluated a small-group tutorial program with a strong phonics emphasis in rural Nacogdoches, Texas. Professors at Stephen F. Austin University served as the tutors, and provided daily 20-minute lessons in groups of 3-4 for 15 weeks. The students were first grade boys who had scored very low on the Texas Primary Reading Inventory. Students were 42% African American, 34% Hispanic, and 20% White. They were randomly assigned to small-group tutorial (n=35) or control (n=29), which did not receive any additional instruction. On spring TPRI tests, adjusted for pretests, there were no significant differences. Effect sizes were -0.05 for Graphophonemic Knowledge, -0.01 for Reading Accuracy, and -0.24 for Reading Comprehension, for a mean of -0.10.

Early Intervention in Reading

Taylor, Short, Frye, & Shearer (1992) developed and evaluated a small group intervention called *Early Intervention in Reading (EIR).* The program focuses on phonemic segmentation and blending, phonics instruction (making words, writing for sounds), story reading and re-reading, comprehension (high level talk and writing about text), and home

reading. In addition to 15-20 minutes of small group teaching, students receive one-to-one or one-to-two tutoring from a paraprofessional 5-10 minutes daily. In a small study in a suburban Midwestern district, 30 struggling first graders were matched with 30 similar students. On Gates MacGinitie posttests, adjusted for pretests, students in *EIR* scored substantially higher (ES=+0.82).

Read, Write, and Type (RWT)-Small Group

Read, Write, and Type is a computer-assisted instruction program that Torgesen, Wagner, Rashotte, Herron, & Lindamood (2009) used to create a small group teaching intervention in which specially trained teachers added to schools' staffs worked with struggling first graders in groups of 3. Children received 4 50-minute sessions per week over the course of the school year for a mean of 80.4 hours. Half of the instructional time involved teacher led instruction, and half involved working on computers to practice skills taught by the teachers. In an evaluation by Torgesen et al. (2009), children within 3 schools were randomly assigned to *RWT* (n=34), to a Lindamood computer-assisted small group intervention (see below), or to an untreated control group (n=39). At posttest at the end of first grade, adjusting for pretests, effect sizes in comparison to controls were +0.41 on Woodcock Word Identification, +0.59 on Woodcock Word Attack, +0.22 on TOWRE Word Efficiency, +0.26 on TOWRE Nonword Efficiency, and +0.33 on Woodcock Passage Comprehension, for a mean of +0.36. A follow-up assessment one

year later of the same measures plus the Gray Oral Reading Test found a mean effect size of +0.22.

Lindamood Phoneme Sequencing Program for Reading, Spelling, and Speech – Small Group (LiPS-Small Group)

The Lindamood Phonics Sequencing Program (LiPS) is a one-to-one tutoring program with a strong phonics focus that teaches children to notice how their mouths make various sounds and relates these to letters and sound blending. *LiPS* is a newer edition of *Auditory Discrimination in Depth (ADD)*, a one-to-one phonetic tutoring program discussed earlier in this article. Torgesen et al. (2009) modified *LiPS* to add computer-assisted instruction and to make it a one to three small group tutorial, *LiPS-Small Group*. As explained in the previous section, Torgesen et al. (2009) randomly assigned first graders within three schools to *LiPS* (n=35), *RWT* (n=34), or control (n=39). On the same measures described for *RWT,* effect sizes for *LiPS* compared to the control group at the end of first grade were +0.63 on Woodcock Word Identification, +0.93 on Woodcock Word Attack, +0.50 on TOWRE Word Efficiency, +0.79 on TOWRE Nonword Efficiency, and +0.46 on Woodcock Passage Comprehension, for a mean of +0.66. On 8 follow-up measures a year later, the mean effect size was +0.40

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Comparisons of Alternative Small Group Models

A few studies compared alternative small group programs. Because these studies lacked control groups representing ordinary practice they do not appear in Table 3, but they are interesting nevertheless.

In a study in southeast Louisiana Reading First schools comparing alternative experimental conditions but lacking a comparison group, Bell (2008) compared struggling third graders who received *Voyager Passport* small group lessons (n=60) with matched students who received *Scholastic Fluency* (n=60) or *Earobics* lessons (n=60). Pretest-adjusted posttest means for each treatment were compared to the average of the other two. On DIBELS Fluency measures, adjusting for pretests, the effect size for *Voyager Passport* students at "some risk" was -0.16 (n.s.), and it was 0.00 (n.s.) for *Voyager* students at "high risk," for a mean of -0.08. For students using *Scholastic Fluency*, the effect size for students "at some risk" was +0.21 (n.s.), and for students "at high risk" it was -0.02 (n.s.), for a mean of +0.10. For students using *Earobics*, the effect size for students at "some risk" was -0.05, and for students at "high risk" it was +0.02, for a mean of -0.02. Although this study could not determine effects of any of the three small group models in comparison to a standard small-group method or to a lack of special interventions, the comparison among the three methods found that *Scholastic Fluency* had somewhat larger positive effect sizes than *Voyager Passport* or *Earobics* for a fluency measure.

Foorman, Francis, Winikates, Mehta, Schatschneider, & Fletcher (1997) carried out a matched evaluation of a small group tutorial program called *Recipe for Reading* in resource

rooms for second and third graders with reading disabilities. The study was intended to compare synthetic and analytic phonics to a sight-word approach, but, in this school-assigned sample, the synthetic phonics group had pretest and IQ scores substantially higher than those of the other two treatments, so only the analytic phonics vs. sight word comparison could be compared. The *Recipe for Reading* program emphasizes an approach in which children work as a whole group to practice onsets and rimes, in which they learned word families (e.g., cat, rat, mat, sat; light, sight, night). Children wrote target words, discussed their meanings, and played games emphasizing how changing initial letters or following letters changed sounds. This program was contrasted to a small group administration of the *Edmark Reading Program*, which emphasizes whole-word identification, matching words and pictures, and using words in increasingly complex sentences. The two approaches were compared in an urban Southwestern school district. There were 46 children in *Recipe for Reading* and 39 matched children in the Edmark sight word program. On posttests, adjusted for pretests, there were no significant differences. On a word reading measure, the effect size was -0.24 (n.s.), favoring the Edmark program.

Hatcher, Goetz, Snowling, Hulme, Gibbs, & Smith (2006) carried out a comparison of two intervention programs for six year olds. One, *Early Literacy Support (ELS),* is widely used in the UK. It provides struggling children in groups of six with 20 hours of structured instruction from a teaching assistant. An alternative method, called *Reading Intervention (RI),* provided 10 hours of group instruction and 10 hours of one-to-one tutoring. The programs also differed in curricula and teaching methods. *ELS* provides a sequence of activities focusing on phonics and

sentence reading using games, songs, and movement as well as direct teaching. In RI, there are 60 20-minute lessons focusing on letter identification, phonological awareness, and writing. Individual sessions included reading an easy book, working on a new book, assessment using running records, and shared reading.

The children were selected from among 16 schools in the North of England. There were 59 *RI* children and 69 *ELS*. Children were pretested, posttested at the end of the 12-week intervention, and then followed up in the next school year, 7 months after the instruction began. Adjusting for pretests, differences immediately after intervention were +0.15 ($p<.05$) favoring *RI* on Letter-Sound Identification, +0.08 (n.s.) on Early Word Reading, and -0.08 (n.s.) on British Abilities Scale Word Reading, for a mean of +0.05. At seven-month followup, there were no significant differences, with effect sizes of +0.09 for Letter-Sound Identification, -0.02 for Early Word Reading, and -0.09 for BAS Word Reading.

Conclusions: Small Group Tutorials

A total of 20 studies evaluated 18 different models of small group tutorials. Sixteen of these used random assignment to conditions. The overall mean effect size was +0.31.

It is important to note that the studies that met the inclusion criteria were not run-of-the-mill Title I pullout programs. Instead, all were named programs with extensive training, materials, and other supports, and a strong emphasis on phonics.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Classroom Instructional Process Programs

One potential solution to the reading problems of many struggling readers is to enhance the quality of instruction in their regular classrooms. In previous reviews of beginning reading and upper-elementary reading programs (Slavin et al., in press), classroom instructional process programs were clearly the most effective and well-evaluated approaches for children in general. Introduction of training in programs that have been successfully evaluated with students in general, such as various forms of cooperative learning and phonics-oriented class programs, could be particularly beneficial for children who would otherwise have difficulty in learning to read. From the perspective of response to intervention, improving classroom instruction is Tier 1 of a plan to improve learning for struggling readers and reduce special education placements. Use of effective classroom strategies does not preclude individually targeted interventions for the hopefully small number of children who still need them, despite high-quality classroom teaching.

Table 4 summarizes research on the outcomes of these types of programs for children in the lowest performance levels of their classes.

==========

TABLE 4 HERE

==========

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Cooperative Integrated Reading and Composition

Cooperative Integrated Reading and Composition, or *CIRC* (Stevens, Madden, Slavin, & Farnish, 1987) is a cooperative learning program in which, following teacher instruction to the whole class, children work in 4-member teams on partner reading, comprehension questions, story-related writing, and comprehension activities. Teams earn recognition based on the average scores earned on individual tests given each week. Several large-scale evaluations have found positive effects of *CIRC* on students' reading gains in grades 2-6 (see Slavin et al., in press). *CIRC i*s used both as a stand-alone program and as the upper-elementary component of the *Success for All* comprehensive reform model, where it is called *Reading Wings* (Slavin & Madden, 2009). This section focuses on effects of this whole-class method on children identified as remedial or special education students.

Stevens & Slavin (1995a) carried out a large 2-year matched evaluation of *CIRC* in a working-class suburb of Baltimore. The experiment included 1299 students in grades 2-6, of which 137 (72E, 65C) were students with learning disabilities. Analyses for these children found effect sizes of +0.40 ($p<.05$) for CAT Vocabulary and +0.31 ($p<.05$) for CAT Comprehension after one year and +0.37 ($p<.05$) for Vocabulary and +0.32 ($p<.05$) for Comprehension after two years. The mean effect size at the end of the study was +0.35.

Stevens & Slavin (1995b) evaluated *CIRC* as part of a larger schoolwide intervention called the *Cooperative Elementary School.* The study took place over a 2-year period in a suburban Maryland district. A total of 1012 students in Grades 2-6 participated. Students were

pretested on CAT reading and then posttested each spring. Separate analyses for students in special education (n=40E, 36C) found effect sizes of +0.29 (n.s.) for Reading Comprehension and +0.26 (n.s.) for Reading Vocabulary after 1 year, but after two years effect sizes were +0.85 ($p<.01$) for Reading Comprehension and +0.76 ($p<.01$) for Reading Vocabulary, for a mean of +0.81.

Bramlett (1994) evaluated CIRC among third graders in schools in rural southern Ohio. 82 students using CIRC were compared to 67 matched control students. Controlling for pretests, CAT posttests slightly favored SFA on Comprehension (ES=+0.33), Total Reading (ES=+0.33), Word Analysis (ES=+0.56), and Vocabulary (ES=+0.30, n.s.), for a mean of +0.38.

Peer-Assisted Literacy Strategies (PALS)

Peer-Assisted Literacy Strategies (PALS) is a technique in which children work in pairs, taking turns as teacher and learner, to learn a structured sequence of literacy skills, such as phonemic awareness, phonics, sound blending, passage reading, and story retelling. Children use a simple error-correction strategy with each other, under guidance from the teacher.

Mathes & Babyak (2001) carried out an evaluation of *PALS* over a 14-week period in a medium-sized district in Florida. Three treatments were compared, but one, a combination of *PALS* and small-group, skills-focused mini lessons, had a duration of less than 12 weeks. The remaining 20 first grade classes were randomly assigned to PALS and control conditions. The three lowest achievers and one average and one high achiever were selected within each class for

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

measurement, so the number of low achieving children was 27 in 10 *PALS* classes and 29 in 10 control classes. The students were 50% White, 48% African American. Adjusting for pretests, effect sizes for PALS compared to control were +0.51 for Woodcock Word Identification, +0.89 for Woodcock Word Attack, +0.71 for Woodcock Basic Skills, and +0.23 for Woodcock Passage Comprehension. Averaging across Woodcock measures, effect sizes were +0.59.

In a 16-week experiment, Mathes, Torgesen, and Allor (2001) evaluated PALS among first graders in a southeastern district. Three treatments were compared, but one, a combination of *PALS* and computerized phonological awareness training, had pretest differences with the control group of more than 50% of a standard deviation. Students were 47% White and 51% African American. Twelve classes were assigned to *PALS* and twelve matched classes were assigned to a control condition. The four lowest achievers and one average and one high achiever were selected within each class for measurement, making the number of low achieving children 42 in *PALS* classes and 33 in control classes. Adjusting for pretests, effect sizes for PALS compared to control were +0.48 for TOWRE Nonword Efficiency, +0.34 for TOWRE Word Efficiency, +0.42 for Woodcock Word Identification, +0.58 for Woodcock Word Attack, +0.55 for Woodcock Basic Skills, and +0.50 for Woodcock Passage Comprehension. Across Woodcock and TOWRE measures, effect sizes averaged +0.48.

Mathes, Howard, Allen, & Fuchs (1998) evaluated PALS in a 16-week study in a southeastern city. Twenty first grade teachers in 6 schools participated. Assignment was partly random and partly matched, so this was considered a matched study. The three lowest achievers

and one average and one high achiever were selected within each class for measurement, so the total sample of low achievers was 28 children in 10 *PALS* classes and 28 children in 10 control classes. *PALS* procedures were used 3 times a week in 35-minute sessions focusing on sounds and words and partner read-alouds, while control classes were described as using traditional whole language models. On Woodcock posttests adjusting for pretests, PALS students outperformed the control students on Word Identification (ES= +0.51), Word Attack (ES=+0.69), and Passage Comprehension (ES=+0.19). Averaging across the 3 scales, the effect size was +0.46.

Mathes, Torgesen, Clancy-Menchetti, Sani, Nicholas, Robinson, & Grek (2003) evaluated PALS with low-achieving first graders in a 16-week study in a southeastern school district. Teachers were assigned to one of three conditions: *PALS* (N = 7 teachers, 31 students); *teacher-directed instruction (TDI),* a small-group model that used the same curriculum but no peer activities (N = 7 teachers, 30 students); and control (N = 8 teachers, 28 students). Although teachers were randomly assigned to *PALS* and *TDI* conditions, some were randomly assigned to the control group while other controls were matched, so the overall design is considered matched. Students in *PALS* classes experienced 35-minute sessions each week, while those in *TDI* received 30-minute sessions each week. The students in the *PALS* condition gained substantially more than controls on all measures, although not all differences were statistically significant. Averaging across five subtests, *PALS* students averaged an effect size of +0.37 in

comparison to controls adjusting for pretests. However, *PALS* students scored nonsignificantly less well than those in the *TDI* condition, adjusting for pretests.

Same-Age Peer Tutoring

Eldredge & Quinn (1988) designed and evaluated a dyadic reading approach with second graders. Children reading below grade level were assigned to pairs with normal-progress reading partners. Grade-level texts were used. The partners "set the pace for reading, read in phase units, and touched each word as it was being read." (p.42). Partners were changed each week. In a year-long experiment classes within 5 middle-class elementary schools in Provo, Utah were randomly assigned to dyadic reading (n=32) or control (n=32) conditions in a randomized quasi-experiment. Control students were taught in homogeneous basal reading groups. Within classes, the lowest achievers within each school's experimental and control classes were matched on pretests and gender. On Gates-MacGinitie posttests adjusting for pretests, the low achievers in the dyad groups scored substantially higher than controls on Comprehension (ES=+1.33, $p<.05$) and Vocabulary (ES=+1.77, $p<.01$) for a mean of +1.55. Put another way, 27 out of 32 students in the dyad group (84%) scored at or above grade level, while 6 of 32 students (19%) scored at grade level in the control group. The authors ascribe the effects to the fact that the reading partners enabled the poor readers to use grade-level texts, while similar students in the control group were taught using below-level textbooks.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading and Integrated Literacy Strategies (RAILS)

Reading and Integrated Literacy Strategies (RAILS) is a professional development approach primarily intended for schools with many children at risk. It provides children in grades K-2 with a 20-minute additional reading period each day to supplement their 60-90 minute regular reading period, and gives teachers extensive professional development focusing on explicit instruction in phonemic awareness, phonics, comprehension, and vocabulary. *RAILS* was evaluated by Stevens et al. (2008) in three matched low-achieving schools in a small city in central Pennsylvania in which 71% of children received free lunch and 94% were White. For students in special education classes in grades K-2 (n=33E, 17C), the MAT effect size was +0.45, and for those in grades 3-5 (n=18E, 19C), the effect size was +0.49.

Contextually-Based Vocabulary Instruction

Nelson & Stage (2007) evaluated a supplementary intervention, *Contextually-Based Vocabulary Instruction,* in which third and fifth graders received instruction in multiple meanings of vocabulary words. The supplementary instruction took place twice weekly for 20-30 minutes. A 3-month study in a Midwestern school district evaluated the approach. Most students (70%) were White, and 24% were Hispanic. Eight third grade and eight fifth grade classes were randomly assigned to vocabulary supplement (n=159) or control (n=149) conditions, making this a randomized quasi-experiment (RQE). All classes used the same *Scott Foresman* basal textbooks. On Gates MacGinitie Comprehension tests, effect sizes for low achievers averaged

+0.60. On Gates Vocabulary, the effect size was +0.23. The overall effect size for low achievers (n=41E, 32C) was +0.41.

Reading Styles

Reading Styles is an intervention in which children with learning disabilities are assessed on a reading style inventory and then given instruction matched to their favored styles. In a study by Lashell (1986), students with learning disabilities in grades 2-6 were given interventions based on their assessed learning styles: phonics/linguistics, Orton-Gillingham, whole word, individualized, language experience, Fernald word-tracing, and Carbo recorded books. Each child received 3-5 different approaches. The treatments were administered during daily one-hour resource room sessions. The study compared all children with learning disabilities in a rural district in Snohomish County, Washington (n=47) to all children with learning disabilities in a similar rural district in the same county (n=43). On the Gray Oral Reading Test, adjusted for pretests, the results significantly favored the *Reading Styles* students (ES=+0.79, $p<.001$).

Brooks (1991) evaluated *Reading Styles* among students in grades 2-6 in northwest Ohio. Students in *Reading Styles* (n=22) were matched with those in a control class (n=20) in a semester-long experiment. On the Spadafore Diagnostic Reading Test, adjusted for pretests, overall effect sizes were +0.21 on Oral Reading and +0.51 for Silent Reading, for a mean of +0.36.

Direct Instruction

Direct Instruction (DI) is a structured, phonetic approach in which teachers use step-by-step materials and methods to help children master decoding and comprehension skills.

Bowers (1972) carried out a small randomized evaluation of *DI* with culturally disadvantaged first graders in eight classes in Oklahoma. Children scoring below the 25th percentile on the Metropolitan Reading Readiness Test (MRRT) were randomly assigned to use *DI* (n=60) or traditional basal texts (n=63). All children were White. Adjusting for pretests, *DI* students scored higher than controls on the Gates McGinitie Comprehension scale (ES=+0.17, p<.05) and the Vocabulary scale (ES=+0.35, p<.05), for a mean effect size of +0.26.

Davis (1995) evaluated *DI* among at-risk students in two schools in southern Mississippi. Teachers of grades 1-3 identified children they considered to be at risk, and those children were put in classes of 10. One school used *DI* and one used traditional basal textbooks. On SAT pretests, there were strong pretest differences for first and third graders, so those results were excluded, but second grade pretests were similar. Adjusting for pretests, SAT Reading posttest differences for second graders (n=59E, 52C) averaged +0.49.

Project Read

Project Read is a phonetic approach to beginning reading instruction based on the Orton-Gillingham method, originally designed for tutoring students with dyslexia. *Project Read* similarly uses a systematic phonics progression, systematic approaches to building

comprehension, writing, and spelling, and extensive professional development for teachers. Greene (1991) carried out an evaluation of *Project Read* in three Louisiana school districts. Children in grades 1-3 who were reading below the 25th percentile in *Project Read* schools were individually matched based on pretests and demographics with control children (n=112E, 112C). On CAT reading, comparisons of gain scores found effect sizes of +0.59. Particularly large relative gains were found for Title I and African American students.

A small study by Acalin (1995) compared children in K-4 special education resource classes taught in *Project Read* (in small groups) to those taught one-to-one using *Reading Recovery*. This study did not qualify for inclusion because it lacked a control group representing ordinary practice, but it is interesting to note that students in *Project Read* scored non-significantly higher than those in *Reading Recovery* on Woodcock Johnson scales (ES=+0.12).

Precision Teaching

Haring & Krug (1975) created and evaluated a precision teaching approach designed to help disadvantaged, mostly African American children with mental retardation to learn to read. Two experimental classes (n=24) were matched with two control classes (n=30). The students were ages 9-12, and the schools were in a large center city. Both groups had average IQs of 72. The experimental classes used structured, phonetic materials, including DISTAR and Sullivan readers, over a full school year. A token economy behavior modification plan was used in which children earned points for correct responses, which they could exchange for toys, activities, and

free time. On WRAT-Reading, controlling for pretests, the posttest effect size was +1.18. Eight of the 24 students qualified for placement in regular classes, while none of 30 control students qualified.

Conclusions: Classroom Instructional Process Programs

The effect sizes across 16 studies of classroom instructional process programs (1 randomized and 3 randomized quasi-experiments) were very positive for students at the lowest performance levels in their classes. The weighted mean was +0.56, similar to the findings for one-to-one phonetic tutoring. These effects are markedly more positive for low achievers than they were for the overall samples of students of all performance levels, reported in Slavin et al. (in press). Eight of the 16 studies involved forms of cooperative learning, *CIRC* (ES=+0.46 in 3 studies), *PALS* (ES=+0.49 in 4 studies), and same-age tutoring (ES= +1.55 in one small study). The mean for all eight studies of cooperative learning was +0.58. Other particularly positive effects were found for programs that utilize structured, systematic, phonetic approaches to reading instruction: *Direct Instruction, Project Read, RAILS,* and *Precision Teaching*.

Classroom Instructional Process with Tutoring (Success for All)

This section presents research on a single program, *Success for All*, which provides extensive school staff training and materials to improve all aspects of school organization and functioning, especially those aspects relating to reading, and also provides tutoring to struggling

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

children, mostly first graders. The classroom interventions use a structured, fast-paced approach with a strong emphasis on cooperative learning, phonics, metacognitive skills, and frequent assessment. In second grade and beyond, *Success for All* uses an adaptation of *CIRC*, described earlier. Parent involvement and interventions for behavior and other non-academic problems are also emphasized. In contrast to one-to-one tutoring programs such as *Reading Recovery*, which provide intensive tutoring in first grade but no intervention afterwards, *Success for All* continues to provide classroom-level interventions (though not tutoring) throughout the elementary grades. Research on the reading outcomes of *Success for All* for students in the lowest performing segments of their classes is summarized in Table 5.

==========

Table 5 Here

==========

The largest and most important evaluation of *Success for All* was a three-year longitudinal cluster randomized experiment (Borman, Slavin, Chamberlain, Madden, & Chambers, 2007). In this study, 35 Title I schools throughout the U.S. were randomly assigned to use *Success for All* either in grades K-2 or 3-5. K-2 children assigned to SFA in one school served as controls for K-2 children not assigned to SFA in another school. A total of 2108 children (1085 T, 1023 C) remained in the study schools all three years, 63% of those originally tested in kindergarten. Attrition was equal in the two treatment groups. Among the final sample,

72% of students received free lunches, and 57% of students were African American, 31% were White, and 10% were Hispanic.

Children were pretested on the PPVT and then individually tested on scales from the Woodcock Reading Mastery Test each spring for three years. Testers were not aware of the treatment assignments of each school. Data were analyzed using HLM, with children nested within schools. For students in the lowest third of their classes on pretests (362E, 341C), effect sizes were +0.22 ($p<.05$) for Word Identification, +0.40 ($p<.01$) for Word Attack, and +0.22 ($p<.05$) for Passage Comprehension, for a mean of +0.28.

The longest evaluation of *Success for All* was a longitudinal matched study of the five original *SFA* schools in Baltimore (Madden, Slavin, Karweit, Dolan, & Wasik, 1993; Slavin, Madden, Dolan & Wasik, 1993). In this study, students in five inner-city Baltimore schools were individually matched with those in similar control schools. Individual matching was based on spring kindergarten CTBS or CAT scores administered by the district, and school matching was based on free lunch and historical achievement levels on district standardized tests. All children were African American, and approximately 95% of children qualified for free lunches.

Each spring, children in all *SFA* and control schools who had begun in their schools by first grade were individually assessed on Woodcock Word Identification, Word Attack, and Passage Comprehension tests. Students in grades 1-3 were also given the Durrell Oral Reading Test, while those in grades 4-5 were given the Gray Oral Reading Test. Testers were not made aware of the schools' treatment assignments. Children were followed and tested as long as they

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

remained in their schools, even if they were retained or assigned to special education. Each year, an additional cohort was added.

A major report on the evaluation was published in the *American Educational Research Journal* after three years (Madden et al., 1993). At that point, the third grade cohort had been in *SFA* or control schools for three years, the second grade for two, and the first grade for one. Averaging across the four measures, the mean pretest-adjusted effect sizes for children whose kindergarten scores put them in the lowest 25% of their grades found significant ($p<.001$) and substantial positive effect sizes for this subgroup: ES=+0.98 for third graders, ES=+1.00 for second graders, and ES=+0.82 for first graders.

Data collected after six years, when the most advanced cohort was in fifth grade, showed substantial positive effects for students who had been in the lowest 25% of their grades at pretest. Averaging across three Woodcock and two Gray scales, effect sizes were +1.03 for fifth graders (34E, 41C), and +0.80 for fourth graders (39E, 41C). Averaging across three Woodcock scales and Durrell Oral Reading, effect sizes were +1.32 for third graders (40E, 48C), +0.92 for second graders (53E, 61C), and +1.18 for first graders (66E, 77C), for a mean of +1.05.

Beyond the achievement effects, Slavin et al. (1993) also reported a substantial difference in retention rates between *SFA* and control schools. By first grade, 34.9% of control students but only 11.2% of *SFA* students had been held back ($p<.001$). According to state data, third grade absences in 1993 were 8.8% in *SFA* schools and 13.5% in control, and among fifth graders the rates were 6.4% in *SFA*, 13.7% in control.

Borman & Hewes (2003) carried out a follow-up assessment of children in the first four Baltimore cohorts when they were in the eighth grade (if they had been promoted each year). Since *SFA* schools only went to the fifth grade, these students would have been out of the *SFA* program for at least 3 years. Effect sizes on CTBS were still significantly positive for the lowest achievers (ES=+0.34). The *SFA* students were also significantly less likely to have been retained or assigned to special education.

Ross, Nunnery, & Smith (1996) evaluated *Success for All* in first grades in two schools in the Amphitheater District near Tucson, Arizona. Each school was matched with two control schools based on prior achievement, percent free lunch, and ethnicity. Overall n's were 169T, 371C. About 23% of children were Spanish-dominant and 13% were ELLs. For students in the lowest quartile at pretest (n=42E, 96C), effect sizes were +0.44 for Woodcock Word Identification, +1.07 for Woodcock Word Attack, +0.30 for Woodcock Passage Comprehension, and +0.37 for Durrell Oral Reading, for a mean of +0.54.

Ross & Casey (1998b) studied *SFA* in 8 schools (151E, 205C) in Ft. Wayne, Indiana that were 75% free lunch and 45% minority (mostly African American). Students were pretested in kindergarten and posttested at the end of first grade. The mean effect sizes across Woodcock and Durrell measures for students in the lowest 25% at pretest (n=36E, 56C) was +0.35.

Ross, Smith, & Casey (1994) evaluated *Success for All* in a White, working class school in Caldwell, Idaho. Students were involved in *SFA* over a period of up to 3 years. They were

compared to a control school. The mean effect size across Woodcock and Durrell measures for low 25% students in Grades 1-3 (n=58E, 38C) was -0.20.

Casey, Smith, & Ross (1994) evaluated *SFA* in three high-poverty African American schools in Memphis. First graders in the schools were matched with those in a single comparison school (n=49E, 16C). Averaging across Woodcock and Durrell measures, adjusted for pretests, the average effect size for the lowest quartile of students was +0.54.

A three-year study in Montgomery, Alabama by Ross, Smith, Bond, Casey, & Johnson (1993) compared two *SFA* and two matched control schools. Students were mostly African American and qualified for free lunch. On Woodcock and Durrell measures, controlling for PPVT, students in first grades who were in the lowest 25% of their classes in kindergarten (n=23E, 24C) had an average effect size on Woodcock and Durrell measures of +1.16.

Ross, Smith, & Casey (1995) carried out a 4-year longitudinal evaluation of *SFA* in four schools in Ft. Wayne, Indiana. Two schools used *SFA* and two were controls. Adjusting for pretests, students in the lowest 25% of their classes at pretest (n=24E, 17C) who remained in the school in grades 2-4 had an average effect size on Woodcock and Durrell measures of +0.45.

Smith, Ross, & Casey (1994) carried out a four-year longitudinal evaluation of *Success for All* in one high-poverty African American school in Memphis. A matched school served as a control. Students in grades 1, 2, and 4 were followed for up to 4 years. Among those in the lowest 25% of their grades (n=21E, 17C), the mean effect size averaging Woodcock and Durrell measures was +1.14.

Conclusions: One-to-One Tutoring with Classroom Instructional Process Approaches

The weighted mean effect size for the lowest achievers in *Success for All* across 9 qualifying studies was +0.52, similar to the effect size for phonetic tutoring programs. Where the results are different, however, is in long-term outcomes. Most of the *SFA* studies provided the program over at least a three-year period, and generally found stable or increasing effect sizes over the years (see, for example, Borman et al., 2007). Slavin et al. (1993) evaluated continued treatment over six years (K-5), and then a followup study by Borman & Hewes (2003) followed *SFA* students to eighth grade. These long-term studies found that positive effects of *Success for All* maintained over time. This is in contrast to the findings of long-term follow-ups of one-to-one tutoring alone without classroom interventions after tutoring. These findings may suggest more broadly that even the most effective first grade tutoring approaches require followup with ongoing intervention in the later grades. This issue is discussed further later in this article.

Instructional Technology

Over the past 30 years, one of the most common solutions applied for children who are struggling to learn to read is to give them computer-assisted instruction (CAI) software. Modern CAI programs adapt to children's specific needs and give them activities with graphics and exciting elements that can supplement classroom instruction. However, previous reviews of research on elementary CAI applications in reading find few positive effects (Dynarski et al.,

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

2007; Kulik, 2003; Slavin et al., in press). Table 6 summarizes research on outcomes of instructional technology for children who are at the lowest performance levels of their classes.

==========

TABLE 6 HERE

==========

Jostens

Becker (1994) evaluated *Jostens* with grade 2-5 students in a high-poverty school in Baltimore. A total of 56 low-achieving students were matched and then randomly assigned to use the *Jostens* integrated learning system in either reading or in math. The *Jostens* group achieved non-significantly better scores on the California Achievement Test than did students who did not use the reading software (ES=+0.41).

Sinkis (1993) evaluated *Jostens* with Title I students in a pullout program in 8 schools in an urban district in the northeast. Four schools used *Jostens* and four served as matched controls. Students in grades 2-6 were involved, but second and fourth grade pretests were more than 50% of a standard deviation apart. Among third graders (n=71E, 63C), MAT Reading Comprehension posttests adjusted for pretests had an effect size of +0.14 (n.s.). Corresponding effect sizes for fifth graders (n=83E, 61C) were +0.22 (n.s.), and for sixth graders (n=74E, 70C) the effect size was -0.01 (n.s.), for a mean across grades of ES=+0.12.

Standish (1995) evaluated *Jostens* among second graders in two suburban Delaware schools. The *Jostens* schools had 4 teachers and 56 students, while the control school had 5 teachers and 83 students. The schools were well matched on cognitive ability tests and demographics. On MAT6 Reading posttests, adjusted for cognitive ability tests and demographic variables, the effect size for a Title I subgroup (n=22E, 21C) was +0.55.

Fast ForWord

Fast ForWord, published by Scientific Learning, is a computerized program designed on the theory that many children with reading and language delays have auditory processing disorders. It uses computer games that slow and magnify acoustic changes within normal speech to "retrain the brain" to process information more effectively. The program was developed by neuroscientists who demonstrated that having children use computer games of this type showed improvements in "temporal processing" skills (Merzenich et al., 1996; Tallal et al., 1996). The initial model was expanded into software for use in schools, adding exercises on reading skills such as word recognition, decoding, fluency, spelling, and vocabulary. Children participate in *Fast ForWord* 90-100 minutes per day, 5 days a week, for 6-8 weeks, so it is intended to make a substantial difference in a relatively short time.

While many studies of *Fast ForWord* have been done, most did not qualify for the current review. Most were too brief (less than 12 weeks), and most used measures of language, not reading. The most rigorous of the brief studies, an 8-week randomized evaluation by

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Borman & Rachuba (2009), found no differences between *Fast ForWord* and control students on reading measures.

The one randomized study of *Fast Forward* that met 12-week duration criterion is an evaluation by Rouse & Krueger (2004). That study involved four schools in a Northeastern city. All schools were implementing *Success for All* (Slavin & Madden, 2009). About 66% of students were Hispanic and 27% were African American, 59% qualified for free or reduced-price lunches, and 61% came from homes in which a language other than English was spoken. Children in grades 3-6 who were in the bottom 20% on the state's standardized test and had parent permission were randomly assigned to the *Fast ForWord* (n=237) or control (n=217) conditions. Students in the *Fast ForWord* group participated in one of two eight-week "flights" in spring, 2001. Students in grades 3 and 5 received an average of 35 days of treatment in January-March, and those in grades 4 and 6 received an average of 28 days in March-June. A variety of measures were given just before and just after treatment, and thus did not meet the duration requirement of 12 weeks. However, the study analyzed state reading test data from Fall, 2000, and Fall, 2001. On posttests adjusted for pretests, there were no differences between *Fast ForWord* and control students (ES=+0.05, n.s.). Sub-analyses of data for children who received the full treatment also showed no differences. Finally, outcomes were near zero on the immediate posttests.

Marion (2004) evaluated *Fast ForWord* in fifth and sixth grades in rural Appalachian Grainger County, Tennessee. Almost all students were White, and 52% received free or

reduced-price lunches. Students who received *Fast ForWord* (N=215) were matched with those who did not (N=134) on Terra Nova pretests. On Terra Nova posttests, adjusted for pretests, *Fast ForWord* students in the lowest quartile (n=34E, 29C) scored non-significantly higher (ES=+0.15, n.s.).

Lexia

Lexia Learning Systems has two supplemental computer-assisted instruction programs: *Phonics Based Reading (PBR)* and *Strategies for Older Students (SOS)*. They consist of various activities that teach phonetic word-attack strategies to promote automaticity in word recognition. Students typically participate in 2 to 4 20-30-minute sessions a week. Macaruso and his colleagues evaluated the *Lexia* programs in a year-long study in 10 first-grade classes in 5 Boston schools (Macaruso et al., 2006). One class in each school was assigned to the experimental group and another to the control group. Over 50% of the 83 students in the experimental group and the 84 students in the control group were eligible for free or reduced-price lunch. Controlling for pretests, the mean effect size for all students was +0.22 on the Gates-MacGinitie Reading Test. For Title 1 students, the effect size was +0.67.

Other Supplemental CAI

Dynarski, Agodini, Heaviside, Novak, Carey, & Campuzano (2007) evaluated the use in first grades of five CAI reading programs, *Destination Reading, Waterford, Headsprout, Plato Focus,* and *Academy of Reading*. Outcomes for individual programs were not reported, so this is

an evaluation of modern uses of technology in first grade reading in general, not of any particular approach. The study involved 43 schools in 11 districts. A total of 158 teachers (89E, 69C) and their 2619 students (1516E, 1103C) were randomly assigned within schools to CAI or control conditions. CAI students used the programs 94 minutes per week, on average. Control classes also often had computers, and used them for purposes such as reading assessment and practice, averaging 18 minutes per week.

Schools involved in the study were very diverse, and were located throughout the US. However, they were relatively disadvantaged, with 49% of students eligible for free or reduced-price lunches and 76% of schools receiving Title I. Overall, 44% of students were White, 31% African American, and 22% Hispanic.

Students were pre- and posttested on the SAT-9. There were no differences for students in general. N's for the lowest 33% of students were 505E, 367C. An analysis of effects on the number of children who had posttests below the 33rd percentile found no differences (ES=+0.02, n.s.).

Campuzano, Dynarski, Agodini, & Rall (2009) reported outcomes for a smaller second cohort of first graders, most of whom were taught by a subset of the same teachers as those in the first cohort. Four of the five programs remained in use, *Destination Reading, Waterford, Headsprout,* and *Plato Focus.* The numbers of first graders in the lowest third of their classes was 130E, 102C. The technology products were used less than half as often in the second yar (19.7 hours) as in the first (42.6 hours). Controlling for pretests, the posttest effect size for the

number of children scoring below the 33rd percentile was -0.39. A weighted average effect size across the two cohorts was

-0.07.

The same study evaluated four CAI programs at the fourth grade level: *Leapfrog, Read 180, Academy of Reading,* and *Knowledge Box*, used an average of 98 minutes per week. Overall, 64% of these students were eligible for free- or reduced price lunches, 57% were African American, 23% were Hispanic, and 17% were White. 118 classrooms (63E, 55C) were randomly assigned to treatments, with 2265 students (1231E, 1034C). N's for the lowest 33% were 410E, 345C. On SAT10, there were no differences in the proportions of students scoring below the 33rd percentile (ES= -0.01).

Campuzano et al. (2009) also reported second-cohort data for fourth graders taught by a subset of the teachers who taught the first cohort. Two of the four first-cohort programs remained in use: *LeapTrack* and *Academy of Reading.* N's were 52E, 43C. The programs were used somewhat more often in the second year (16 hours) than in the first (12 hours). Effects on the number of children scoring below the 33rd percentile were nonsignificantly positive (ES=+0.48). A weighted average effect size for the two cohorts was +0.04.

Becker (1994) reported a randomized evaluation of an ILS program called CNS. A total of 60 low-achieving students in grades 2-5 in an integrated Baltimore school with 50% of children receiving free lunch were randomly assigned within 9 classes to use CNS either in

reading or in math. The math students served as a control group in the reading evaluation. On CAT reading scores controlling for pretests, effect sizes for low achievers averaged +0.10 (n.s).

Ramey (1991) carried out an evaluation of several interventions for low-achieving students in Seattle. One of these was computer-assisted instruction in reading. Matched students in grades 2-5 in 1989-90 received either CAI (n=62) or traditional pullout instruction in small groups (n=220). (There was also an untreated control group, but its pretest scores were too high to qualify in this review.) On CAT-Reading, adjusted for pretests, effect sizes were +0.22 (n.s.) at the end of the treatment year and +0.24 (n.s.) at the end of a follow-up year.

In a small study in two Virginia Title I schools, Bass, Ries, & Sharpe (1986) evaluated the use of a variety of software in grades 5-6. Students in one school using CAI (n=73) were compared to those in a matched school (n=72). Students were pre- and posttested on the SRA and the Virginia Basic Learning Skills Test. Averaging fifth and sixth grade scores, effect sizes were +0.22 for the SRA and +0.13 for the BLS, for a mean of +0.18.

Chiang et al. (1978) evaluated a supplementary CAI strategy with students with a variety of special needs, mainly learning disabilities. The students ranged in age from 7 to 12, with an average age of 10. They were in four schools in Cupertino, California, a middle class suburb of San Francisco. Students using CAI (n=65) were individually matched with similar students in other schools in the district who did not use CAI (n=72). On PIAT Reading Recognition (given in May), controlling for September pretests, the effect size was +0.18 (n.s.), and it was +0.26 (n.s.) for PIAT Reading Comprehension.

Roth & Beck (1987) carried out an evaluation of two microcomputer programs, called *Construct-a-Word* and *Hint and Hunt*, designed to build decoding and word recognition skills. They compared three fourth-grade classes in a low SES, low achieving urban school to three classes in a matched comparison school. Experimental students used the computers about 20 minutes a day, three times a week, over a school year, in addition to their usual reading instruction. Among low achievers (reading below a grade equivalent of 3.0), n's were 20E, 17C. On CAT Vocabulary, effect sizes were estimated at +0.98, but on CAT Reading Comprehension there were no differences (estimated ES = -0.10). Averaging across the two measures gives a mean ES of +0.44.

Coomes (1985) evaluated the use of a variety of drill and practice software in four middle class schools in Texas. For low achievers (n=18E, 18C), the effect size was non-significant but positive (+0.30, n.s.).

Conclusions: Instructional Technology

Across 14 qualifying studies (5 randomized), IT had minimal impacts on the achievement of struggling readers. The weighted mean effect size was only +0.09.

Findings on Key Issues

The review of programs for struggling readers examined many important questions beyond the overall outcomes of various approaches. The following sections address these questions. In each case, we consider both within-study and between-study comparisons. The

within-study comparisons use the same study inclusion criteria as those applied for the main review, but in a few cases studies that compared alternative treatments but did not qualify for inclusion due to the lack of a control group representing ordinary practice are cited if they met all other inclusion criteria.

The Importance of Phonics

Across all categories of programs, almost all successful programs have a strong emphasis on phonics. As noted earlier, one-to-one tutoring programs in which teachers were the tutors had a much more positive weighted mean effect size if they had a strong phonetic emphasis (mean ES= +0.69 in 9 studies). One-to-one tutoring programs with less of an emphasis on phonics, specifically *Reading Recovery* and *TEACH*, had a weighted mean effect size of +0.23. (*Reading Recovery* now has more of an emphasis on phonics, but only the Burroughs-Lange (2008) London study took place recently enough to reflect this change). Within-study comparisons support the same conclusion. Iverson & Tunmer (1993) compared *Reading Recovery* as ordinarily used at the time to a version with a phonetic emphasis, and found non-significantly more positive effects for the phonetic version (ES=+0.23). Hatcher et al. (1994) also compared a *Reading Recovery*-like treatment to the same treatment with a strong phonology component and to a tutoring model focusing only on phonology (which did not involve reading of real books). The combination of phonology and reading was much more effective than the *Reading Recovery*-like treatment (mean ES=+0.39). Averaging ordinary *Reading Recovery* and a version

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

with less training, a study by Pinnell et al. (1994) found few differences between these programs and a phonetic tutoring model (ES=+0.07 favoring *Reading Recovery*), and a comparison of *TEACH* to a phonetic model also found few differences (ES=+0.07 favoring *TEACH*). Hurry and Silva (2007) compared *Reading Recovery t*o a phonetic tutoring model and found strong differences in favor of *Reading Recovery* at age 7 (ES=+0.73), but a followup to age 10 found that children who had had one-to-one phonology training scored slightly higher than those who had had *Reading Recovery*. Averaging across all comparisons of within-study comparisons (n=5), the mean difference was +0.18 favoring approaches with a phonology emphasis, with the most positive effects for combinations of real reading with systematic phonics.

Long-Term Effects of First Grade Tutoring

A key part of the argument for intensive one-to-one tutoring for struggling first graders depends on the idea that tutoring can bring most struggling readers up to grade level and then they will remain normal readers from then on. The cost-effectiveness argument for very expensive programs such as *Reading Recovery* make the case that the expense is justified by reductions in the need for remedial services or retentions in later years.

Surprisingly, few long-term follow-ups of outcomes of early tutoring have been done, but the existing evidence does not support the contention that successful first grade tutoring has long-lasting effects. The best study of these long-term outcomes was reported by Hurry & Sylva (2007), who followed up London six and seven year olds who received *Reading Recovery*. At the

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

end of their tutoring year, the *Reading Recovery*-tutored children scored substantially better than matched children who were not tutored (ES=+0.85, p<.001). However, a year later, at age 8, the effect size dropped to +0.40 (p<.01), and then to +0.15 (n.s.) at age 10. These data are summarized in

Figure 1.

==========

FIGURE 1 HERE

==========

The Hurry & Sylva (2007) study is most interesting because it did find a strong effect on treatment-independent measures initially, but they dissipated over time. Pinnell, DeFord, & Lyons (1988) found a similar pattern on *Reading Recovery*'s treatment-inherent Text Reading Level measure. Effect sizes were +0.80 in first grade, +0.47 in second grade, and +0.26 in third grade. Another 5-year followup study in Detroit by Huggins (1999) found no differences on state accountability tests between former *Reading Recovery* and control students in fifth grade (ES=+0.13, n.s.), but this study also found no differences at first grade (ES= -0.09, n.s.). A followup study of *Reading Recovery*-tutored students in North Carolina by Baenen, Bernholc, Dulaney, & Banks (1997) found no differences among third graders who had been tutored in first grade with *Reading Recovery* and matched control students on North Carolina End-of-Grade tests. The control group scored slightly higher than former *Reading Recovery* students on the state tests, and there were no differences in retentions, special education placements, or

qualification for Title I remedial services. The *Reading Recovery* students had scored higher than controls on Clay Diagnostic Survey measures at the end of first grade, but there were no treatment-independent measures at that point in time.

While there is little evidence to suggest that first grade tutoring alone is sufficient to maintain struggling students at high levels, an interesting point of contrast comes from a long-term Baltimore evaluation of *Success for All*. Students in the lowest 25% of their schools were followed by Slavin et al. (1993) from first to fifth grades, and were tested each year on Woodcock and Durrell measures. The outcomes are summarized in Figure 2.

==========

FIGURE 2 HERE

==========

As is clear from Figure 2, the initially lowest-achieving students in *Success for All* scored substantially better than matched controls at the end of first grade (ES=+1.18, $p<.001$), and they maintained these gains through fifth grade (ES=+1.03, $p<.001$). A follow-up to eighth grade by Borman & Hewes (2003) found that the former *SFA* students continued to score significantly better than controls on CTBS Reading (ES=+0.34, $p<.01$), and were significantly less likely to have been retained or assigned to special education. Numerous three- and four-year longitudinal studies have also found continuing positive effects of Success for All over time (e.g., Borman et al., 2007, Correnti, 2009; Ross, Smith, & Casey, 1995; Ross et al., 1993). The comparison of

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

these longitudinal studies suggests that tutoring in first grade can have substantial and lasting effects if the tutoring is followed up with improvements in classroom instruction throughout elementary school.

Teachers vs. Paraprofessionals and Volunteers as Tutors

The cost of having certified teachers tutor struggling readers is prohibitive for many schools, so tutoring is often provided by paraprofessionals. Volunteers are also often used as tutors, and volunteer tutoring was the focus of the Clinton Administration's America Reads initiative. Can paraprofessionals and volunteers do as well as teachers in tutoring struggling readers?

The outcomes summarized in Table 2 suggest that paraprofessionals and volunteers can obtain good results with struggling readers, with a mean effect size across 18 studies of +0.24. The mean effect size for paraprofessionals was +0.38 in 11 studies, and for volunteers it was +0.16 in 7 studies (but excluding two studies in which volunteers tutored only once or twice a week, the mean was +0.50). The overall effects for paraprofessionals are similar to those found for all studies of one-to-one tutoring by teachers (ES=+0.38). However, almost all of the paraprofessional tutoring studies involved programs with a strong phonetic emphasis, so a better comparison is between phonetic tutoring by teachers (ES=+0.69) and phonetic tutoring by paraprofessionals (ES=+0.38). One small study, by Brown et al. (2005), compared teachers and paraprofessionals as tutors using the same program, and found much better effects for teachers

(ES=+0.47). Still, children tutored by paraprofessionals scored much better than similar children who were not tutored (ES=+0.53). Ehri et al. (2007) found much better outcomes for teachers than for paraprofessionals using the same program (ES=+0.52), but again, students tutored by paraprofessionals obtained much better outcomes than non-tutored controls (ES= +0.89).

One-to-One vs. Small Group Tutorials

For many years, the dominant form of assistance for struggling readers has been small group tutorials, in which students are taught in groups of two or more. This has been the main use of Title I/Chapter 1 funds since Title I began in 1965, and was emphasized in the Bush Administration's Reading First and Supplemental Educational Services programs. Current policies on response to intervention (RTI) also suggest small-group tutorials as Tier 2 interventions for struggling readers (see Gersten et al, 2009).

The renewed emphasis on small-group tutorials in US policies was motivated in part by a review of research by Elbaum et al. (2000) that concluded that small group and one-to-one instruction had similar effects. However, this conclusion was based on just two small dissertations. One, by Evans (1996), did not meet the inclusion requirements as it involved only 8 students within a single class and used only Clay Diagnostic Survey measures. The other, by Acalin (1995), compared *Reading Recovery* to *Project Read,* a small-group approach that provides extensive training in phonics and phonemic awareness, hardly a run-of-the mill small group treatment.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

The data summarized in Table 3 suggest that small group tutorials with a strong phonetic emphasis and extensive training and followup can indeed be effective for struggling readers (weighted mean ES=+0.31 in 20 studies), but the effects are less than those for phonetic 1-1 tutoring by teachers (mean ES=+0.69) and similar to the effect size for 1-1 tutoring by paraprofessionals (ES=+0.38). Within-study comparisons by Ehri et al. (2007) found much larger effects for one-to-one tutoring than for a small-group treatment that used a similar curriculum (ES=+0.57).

A study by Vaughn, Linan-Thompson, Kouzekanani, Bryant, Dickson, & Blozis (2003) compared 1-1, 1-3, and 1-10 groupings for struggling second graders. Instructional methods were held constant across groupings. Averaging Woodcock Word Attack, Woodcock Passage Comprehension, and DIBELS Fluency, adjusted posttest effect sizes for monolingual English speakers were +0.32 for 1-1 vs. 1-3 and +0.71 for 1-1 vs. 1-10. For English language learners there were no differences between 1-1 and 1-3 (ES=+0.06), but the effect size for 1-1 vs. 1-10 was +0.43.

Simply multiplying the effect sizes for small group teaching by the number of students involved would suggest that small group tutoring is more cost-effective, although it is important to note that most small group tutoring programs provide 30-40 minutes of daily instruction all year, while one-to-one tutoring is usually given less time per day and for a few months, so differences in teacher time per child are not as large as they might appear. For example, a teacher could equally teach a group of three all year or provide 60 tutoring sessions to each of

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

three children. These findings are important in considering response to intervention, for example, because it suggests that if they can afford it, schools should try to arrange one-to-one tutoring for students in the greatest difficulty. There may also be a broader set of students with milder difficulties for whom small group tutorials are sufficient. The tradeoff between the number of students served and the effectiveness of tutoring needs to be carefully considered.

Classroom Instructional Process Approaches vs. Tutoring

One of the most surprising findings in the present review is the effectiveness of classroom instructional process approaches. The average effect size across 16 studies was +0.56 for students in the lowest performing segment of their classes, similar to the effect sizes found for one-to-one phonetic tutoring. Most of these instructional process programs are forms of cooperative learning with a strong focus on phonics (e.g., *CIRC* and *PALS*), and other structured phonetic models (e.g., *Direct Instruction, RAILS,* and *Project Read*).

What these findings suggest is that it is critical to focus first on core classroom instructional strategies, using methods for the whole class that improve reading performance for all, but particularly for low achievers. These same instructional process programs also work with students in general, but the effect sizes for lowest-achieving students are about twice those reported for students in general (see Slavin et al., in press). The finding that improvements in classroom instruction are associated with effect sizes like those of tutoring does not imply that tutoring is unnecessary, but rather that professional development for classroom teachers in

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

proven methods should be a major emphasis of programming for struggling readers. There will always be individual children who continue to struggle despite excellent classroom instruction, but the numbers should be much smaller and the remaining difficulties more tractable when initial classroom instruction has used effective instructional process approaches.

Implications for Response to Intervention

Response to intervention, or *RTI* (Allington & Walmsley, 2007; Fuchs & Fuchs, 2006; Gersten et al., 2009) is a policy, currently dominant in the U.S., that emphasizes the need to provide struggling students with increasing levels of support to attempt to solve their learning problems in the context of general education, before involving the special education system. *RTI* describes three "tiers" of intervention. Tier 1 is regular classroom instruction, Tier 2 is almost always small-group instruction, and Tier 3 may be one-to-one tutoring, other intensive services, or possibly assignment to special education.

The findings of the current review have important implications for *RTI*. First, they emphasize the potential of Tier 1 instruction to enhance the learning of at-risk students. If teachers can significantly enhance the learning of low achievers by adopting cooperative learning or structured phonetic classroom models, this avoids a great deal of frustration, demotivation, and possibly stigmatization for the children themselves, as well as greatly reducing difficulties, expense, and disruption inherent to providing supplemental small-group or tutoring services.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Secondly, the conclusions of this review differ from those of Elbaum et al. (2000), who reported that there were few differences between one-to-one tutoring and small-group tutorials. More than a decade later, there are now many more studies of both types of intervention as well as more within-study comparisons, which show clear differences between phonetic tutoring and phonetic small group instruction. One-to-one, not small group instruction, appears to be the most effective Tier 2 intervention. Because one-to-one tutoring by teachers is expensive, there may be a rationale for Tier 2 interventions that provide students with mild reading problems specific small-group tutoring with extensive professional development, such as *QuickReads, Corrective Reading, Empower Reading, Voyager Passport, Early Intervention in Reading, LiPS, Read, Write, and Type,* or *Read Naturally*. Certainly, the findings of this review would suggest that children who have failed to respond adequately to Tier 1 instruction using proven models and then have failed to respond adequately to proven small-group tutorials should receive one-to-one tutoring using proven models before long-term special education services are considered. The evidence does not support the idea that a relatively brief tutoring experience in first grade is enough to ensure success throughout elementary school and beyond, but it does suggest that with a continuing focus on effective classroom instructional models, most children who receive effecting tutoring interventions in first grade can be kept on track in reading.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Summarizing Evidence of Effectiveness for Current Programs

For many audiences, it is useful to have summaries of the strength of the evidence supporting achievement effects for programs educators might select to improve student outcomes. Slavin (2008) proposed a rating system intended to balance methodological quality, weighted mean effect sizes, sample sizes, and other factors, and this system was applied in all previous best-evidence syntheses. Using similar procedures, programs for struggling readers were categorized as follows.

Strong Evidence of Effectiveness

At least two studies, one of which is a randomized or randomized quasi-experimental study, or multiple smaller studies, with a sample size-weighted effect size of at least +0.20, and a collective sample size across all studies of at least 250 students. To qualify for this category, effect sizes from the randomized studies must have a weighted mean effect size of at least +0.20.

Moderate Evidence of Effectiveness

At least two matched prospective studies, with a collective sample size of 250 students, and a weighted mean effect size of at least +0.20.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

◔ Limited Evidence of Effectiveness: Strong Evidence of Modest Effects

Studies meet the criteria for "Moderate Evidence of Effectiveness" except that the weighted mean effect size is +0.10 to +0.19.

◔ Limited Evidence of Effectiveness: Weak Evidence with Notable Effects

A weighted mean effect size of at least +0.20 based on one or more qualifying studies of any qualifying design insufficient in number or sample size to meet the criteria for "Moderate Evidence of Effectiveness."

○ Insufficient Evidence of Effectiveness

One or more qualifying studies not meeting the criteria for "Limited Evidence of Effectiveness."

N No Qualifying Studies

Table 7 summarizes currently available programs falling into each of these categories.

==========

Table 7 Here

==========

In the category of "strong evidence of effectiveness" were several programs. Success for All, with an effect size of +0.52 in 9 studies, had more evidence of strong effects than any other

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

program. Direct Instruction, a whole-class instructional process approach (ES=+0.37 in 2 small studies) and Corrective Reading, a remedial small group form of Direct Instruction (ES=+0.71 in 2 studies) were considered together as having strong evidence (ES=+.56 in 4 studies). Also in this category were one classroom instructional process approach, Peer Assisted Learning Strategies (PALS) (ES=+0.49 in 4 studies), two one-to-one teacher tutoring programs, Reading Recovery (ES=+0.23 in 8 studies) and Targeted Reading Intervention (ES=+0.29 in 2 studies), and an additional small group tutorial, QuickReads (ES=+0.21 in 2 studies).

There were several one-to-one tutoring programs with phonetic emphases that each had one or more qualifying studies with a large positive effect size, yet all but one of these programs, *Targeted Reading Intervention,* had studies with insufficient sample sizes to qualify for the "strong evidence" category individually. These similar approaches, *Auditory Discrimination in Depth, Early Steps/Howard Street Tutoring, Reading with Phonology, Reading Rescue,* and *Intensive Remediation,* were considered collectively as having strong evidence of effectiveness. Similarly, a set of phonetically-focused tutoring programs delivered by paraprofessionals and volunteers were collectively considered to have strong evidence of effectiveness, although none of them had sufficient study sample sizes to qualify on its own. These were *Sound Partners, The Reading Connection, SMART, Reading Rescue* (with paraprofessionals), *Howard Street Tutoring* (with paraprofessionals), and *Book Buddies*.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

One classroom cooperative learning program was rated as having moderate evidence of effectiveness. This was *Cooperative Integrated Reading and Composition (CIRC),* with an effect size of +0.46 in 3 studies.

The *Jostens* CAI program met the criteria for strong evidence of modest effects, with an effect size of +0.19 in 3 studies. Several programs had a single study with promising effect sizes of +0.20 or more, and were rated as having "weak evidence of notable effects." These are listed in alphabetical order in Table 7. There was insufficient evidence or no evidence at all for a very large number of programs, also listed in Table 7.

Discussion

A total of 96 studies met the inclusion criteria for this review. They compared alternative strategies for helping struggling students in the elementary grades to succeed in reading. Remarkably, 39 of these studies used random assignment to treatments, and five used randomized quasi-experiments. Collectively, the studies involved more than 14,000 children.

Key findings of the review were as follows.

1. One-to-one tutoring works. Teachers are more effective as tutors than paraprofessionals or volunteers, and an emphasis on phonics greatly improves tutoring outcomes.
2. Although one-to-one phonetic tutoring for first graders is highly effective, effects last into the upper elementary grades only if classroom interventions continue past first grade.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

3. Small group tutorials can be effective, but are not as effective as one-to-one instruction by teachers or paraprofessionals.
4. Classroom instructional process approaches, especially cooperative learning and structured phonetic models, have strong effects for low achievers (as well as other students).
5. Traditional computer-assisted instruction programs have little impact on reading.

These findings support the idea, central to current response-to-intervention models (Fuchs & Fuchs, 2006; Gersten et al., 2009), that the best approach for struggling readers is to provide high-quality instruction in the first place, followed up with intensive instruction to the hopefully small number of students who continue to have difficulties despite high-quality classroom instruction. However, the findings point to a particular focus on Tier 1 (regular classroom teaching), and on the use of one-to-one rather than small group instruction for students with the most serious difficulties. The findings support the idea that high quality intervention over many years is needed for lasting impacts, in contrast to the expectation that brief, intensive tutoring will put struggling readers permanently on track. Finally, the findings are consistent with those of previous reviews of classroom instructional programs for elementary reading by Slavin et al. (in press), which found more positive effects of programs that provide extensive professional development to teachers in proven models than they did for programs that provide technology, alternative curricula, or other interventions that do not change daily teaching practices.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

The message of this review is optimistic. There are many proven and promising approaches for struggling readers. It is no longer possible or responsible to do less than what we know how to do to be able to greatly reduce the numbers of children who fail to learn to read adequately. We have both effective and cost-effective tools at hand. While more research is always needed, we already know enough to make a substantial difference in the reading performance of at-risk children.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

References

Acalin, T. (1995). *A comparison of Reading Recovery to Project READ*. M.S. dissertation, California State University, Fullerton, United States -- California. Retrieved September 5, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 1361908).

Allington, R.A., & Walmsey, S. (Eds.) (2007). *No quick fix, the RTI edition: Rethinking literacy programs in America's elementary schools*. New York: Teacher's College Press.

Allor, J. & McCathren, R. (2004). The efficacy of an early literacy tutoring program implemented by college students. *Learning Disabilities Research & Practice, 19(2), 116-129.*

Baker, S., Gersten, R., & Keating, T. (2000). When less may be more: A 2-year longitudinal evaluation of a volunteer tutoring program requiring minimal training. *Reading Research Quarterly, 35*(4), 494-515.

Bass, G., Ries, R. & Sharpe, W. (1986). Teaching basic skills through microcomputer assisted instruction. *Journal of Educational Computing Research, 2*(2), 207-219.

Becker, H.J. (1994). Mindless or mindful use of integrated learning systems. *International Journal of Educational Research, 21*(1), 65-79.

Blachman, B. A., Schatschneider, C., Fletcher, J.M., Francis, D.J., Clonan, S.M., Shaywitz, & B.A., Shaywitz, S.E. (2004). Effects of intensive reading remediation for second and third grades and a 1-year follow-up. *Journal of Educational Psychology, 96* (3), 444-461.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Borman, G., & Hewes, G. (2003*).* Long-term effects and cost effectiveness of Success for All. *Educational Evaluation and Policy Analysis, 24* (2), 243-266.

Borman, G. D., & Rachuba, L. T. (2001). *Evaluation of the Scientific Learning Corporation's Fast ForWord computer-based training program in the Baltimore City Public Schools.* Report prepared for the Abell Foundation, Baltimore, MD.

Borman, G.D., Slavin, R.E., Cheung, A., Chamberlain, A., Madden, N.A., & Chambers, B. (2007). Final reading outcomes of the national randomized field trial of Success for All. *American Educational Research Journal, 44* (3), 701-731.

Bowers, W.M. (1972). *An evaluation of a pilot program in reading for culturally disadvantaged first grade students*. Unpublished doctoral dissertation, University of Tulsa.

Bramlett, R. (1994). Implementing cooperative learning: A field study evaluating issues for school-based consultants. *Journal of School Psychology, 32* (1), 67-84.

Brooks, G. (2007). *What works for children with literacy difficulties: The effectiveness of intervention schemes.* 3rd Edition. DfCSF/NFER: Slough, England.

Brooks, J. D. (1991*). Teaching to identified learning styles: The effects upon oral and silent reading and listening comprehension*. Doctoral Dissertation, University of Toledo.

Brown, K. , Morris, D., & Fields, M. (2005). Intervention after grade 1: Serving increased numbers of struggling readers effectively. *Journal of Literacy Research, 37(*1), 61-94.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Brown, I.S., & Felton, R.H. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. *Reading and Writing. An Interdisciplinary Journal, 2*, 223-241.

Burroughs-Lange, S. & Douetil, J. (2007). Literacy progress of young children from Poor urban settings: A Reading Recovery comparison study. *Literacy Teaching and Learning,* 12 (1), 19-46.

Burroughs-Lange, S. (2007). *Evaluation of Reading Recovery in London Schools: Every Child A Reader 2005-2006*, University of London Department for Education and Skills (2007) *Department for Education and Skills Departmental Report 2007* [online], TSO (The Stationery Office) Available from: http://www.dcsf.gov.uk/aboutus/reports/pdfs/deptreport2007.pdf.

Burroughs-Lange, S. (2008). *Comparison of literacy progress of young children in London schools: A Reading Recovery follow-up study*. Retrieved from http://ioewebserver.ioe.ac.uk/ioe/cms/get.asp?cid=9263.

Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). *Effectiveness of reading and mathematics software products: Findings from two student cohorts*. Washington, DC: Institute of Education Sciences.

Casey, J., Smith., L., & Ross, S. (1994). *Final report: 1993-94 Success for All program in Memphis, Tennessee: Formative evaluation of new SFA schools*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.

Center, Y., Wheldall, K., Freeman, L., Outhred, L. et al. (1995). An evaluation of Reading Recovery. *Reading Research Quarterly, 30*(2), 240-263.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Chiang, A., et al. (1978). *Demonstration of the use of computer-assisted instruction with handicapped children. Final report*. (ERIC Document Reproduction Service No. ED166913)

Cheung, A., & Slavin, R.E. (2005). Effective reading programs for English language learners and other language minority students. *Bilingual Research Journal, 29* (2), 241-267.

Clark, D. E., Jr. (2007) *A comparison of 3rd grade reading scores between students using an integrated literature based curriculum and students using direct instruction at a charter school*. Unpublished doctoral dissertation, Wilmington College.

Clay, M. (1985). *The Early Detection of Reading Difficulties. Third Edition*. (ERIC Document Reproduction Service No. ED263529).

Compton, G. (1992). *The Reading Connection: A leadership initiative designed to change the delivery of educational services to at-risk children*. Unpublished doctoral dissertation, Western Michigan University.

Coomes, P. (1985). *The effects of computer assisted instruction on the development of reading and language skills.* Unpublished doctoral dissertation, North Texas State University.

Cooper, H. (1998). *Synthesizing research (3rd ed.).* Thousand Oaks, CA: Sage.

Correnti, R. (2009, March). *Examining CSR program effects on student achievement: Causal explanation through examination of implementation rates and student mobility.* Paper

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

presented at the annual meetings of the Society for Research on Educational Effectiveness, Crystal City, VA.

Curry, J., Griffith, J., & Williams, H. (1995). *Reading Recovery in AISD*. Austin Independent School District, Department of Audit and Evaluation.

Dorval, B., Wallach, L., & Wallach, M. A. (1978). Field evaluation of a tutorial reading program emphasizing phoneme identification skills. *The Reading Teacher, 31,* 784-790.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., Means, B., Murphy, R., Penuel, W., Javitz, H., Emery, D., & Sussex, W. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort*. Washington, DC: Institute of Education Sciences.

Ehri, L., Dreyer, L., Flugman, B., & Gross, Al. (2007). Reading Rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal, 44*(2), 414-448.

Elbaum, B.,Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology, 92,* 605-619.

Escamilla, K. (1997). Descrubriendo la lectura: An early intervention literacy program in Spanish. In S. Swartz & A. Klein (Eds.), *Research in Reading Recovery* (pp. 109-121). Portsmouth, NH: Heinemann.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Evans, T.L.P. (1996). *I can read deze books: A qualitative comparison of the Reading Recovery Program and a small-group reading intervention*. Unpublished doctoral dissertation, Auburn University.

Foorman, B.R., Francis, D.J., Winikates, D., Mehta, P., Schatschneider, C., & Fletcher, J.M. (1997). Early interventions for children with reading disabilities. *Scientific Studies of Reading, 1*, 255-276.

Fuchs, D., & Fuchs, L.S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41* (1), 92-128.

Gersten, R., Compton, D., Connor, C., Cimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. (2009). *Assisting students struggling with reading: Response to intervention and multi-tier intervention in the primary grades. A practice guide.* (NCEE 2009-4045). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides

Gottshall, D.L. (2007). *Gottshall Early Reading Intervention: A phonics based approach to enhance the achievement of low performing, rural, first grade boys.* Unpublished doctoral dissertation, University of North Texas.

Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education, 34*(2), 90-103.

Gunn, B., Smolkowski, K., Biglan, A., & Black, C. (2002). Supplemental instruction in decoding skills for Hispanic and Non-Hispanic students in early elementary school: A follow up. *Journal of Special Education, 36* (2), 69-79.

Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction: Results for Hispanic and non-Hispanic students. *Journal of Special Education, 39* (2), 66-85.

Hatcher, P.J., Goetz, K., Snowling, M.J., Hulme, C., Gibbs, S., & Smith, G. (2006). Evidence of effectiveness of the Early Literacy Support programme. *British Journal of Educational psychology, 76*, 351-367.

Hatcher, P.J., Hulme, C., & Ellis, A.W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65*, 41-57.

Heistad, D. (2005). *A Minneapolis study of the effects of Read Naturally on fluency and reading comprehension: A supplemental service intervention*. Minnesota: Minneapolis Public Schools.

Hiebert, E. (1994). Reading Recovery in the United States: What difference does it make to an age cohort? *Educational Researcher*, 23(9), 15-25.

Huggins, R. (1999). *Longitudinal study of the Reading Recovery program, 1994-1998*. Detroit Public Schools: Office of Research, Evaluation and Assessment. April 1999.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Hurry, J., & Sylva, K. (2007). Long-term outcomes of early reading intervention. *Journal of Research in Reading 30*(3), 227–248.

Iverson, S. & Tunmer, W. (1993). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology, 85* (1), 112-126.

Jenkins, J. R., Peyton, J. A., Sanders, E. A., & Vadasy, P. F. (2004). Effects of reading decodable texts in supplemental first–grade tutoring. *Scientific Studies of Reading, 8,* 53–85.

Jenkins, J., Jewell, M., Leicester, N., O'Connor, R., Jenkins, L. & Troutner, N. (1994). Accommodations for individual differences without classroom groups: An experiment in school restructuring. *Exceptional Children, 60*(4), 344-358.

Kulik, J. A. (2003). *Effects of using instructional technology in elementary and secondary schools: What controlled evaluation studies say. SRI Project Number P10446.001.* Arlington, VA: SRI International.

LaShell, L. (1986). *An analysis of the effects of reading methods upon reading achievement and locus of control when individual reading style is matched for learning-disabled students*. Doctoral Dissertation, Fielding University.

Lesnick, J. (2006). *A mixed-method multi-level randomized evaluation of the implementation and impact of an audio-assisted reading program for struggling readers.* Unpublished doctoral dissertation, University of Pennsylvania.

Lindamood, C. & Lindamood, P. (1984). *Auditory Discrimination in Depth*. Austin, TX: Pro-Ed.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lovett, M.W., Lacerenza, L., Borden, S., Frijters, J., Steinbach, K., & De Palma, M. (2000). Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes. *Journal of Educational Psychology 92*:2, 263

Macaruso, P., Hook, P.E., & McCabe, R. (2006). The efficacy of computer-based supplementary phonics programs for advancing reading skills in at-risk elementary students. *Journal of Research in Reading, 29*, 162-172.

Madden, N.A., Slavin, R.E., Karweit, N.L., Dolan, L.J., & Wasik, B.A. (1993). Success for All: Longitudinal effects of a restructuring program for inner-city elementary schools. *American Educational Research Journal, 30*, 123-148.

Mantzicopoulos, P., Morrison, D., Stone, E., & Setrakian, W. (1992). Use of the search/teach tutoring approach with middle-class students at risk for reading failure. *The Elementary School Journal, 92*, 573-586.

Marion, G.G. (2004). *An examination of the relationship between students' use of the Fast ForWord Reading Program and their performance on standardized assessments in elementary schools*. Unpublished doctoral dissertation, East Tennessee State University.

Mathes, P.G., & Babyak, A. (2001). The effects of Peer-Assisted Literacy Strategies for first-grade readers with and without additional mini-skills lessons. *Learning Disabilities Research & Practice, 16* (1), 28-44.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Mathes, P.G., Denton, C.A., Fletcher, J.M., Anthony, J.L. Francis, D.J., Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly, 40*(2), 148-182.

Mathes, P.G., Howard, J.K., Allen, S.H., & Fuchs, D. (l998). Peer-Assisted Learning Strategies for first-grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly, 33*, 62-94.

Mathes, P.G., Torgesen, J., Clancy-Menchetti, J., Santi, K., Nicholas, K., Robinson, C., & Grek, M. (2003). A comparison of teacher-directed versus peer-assisted instruction to struggling first-grade readers. *The Elementary School Journal, 103*(5), 459-479.

Mathes, P.G., Torgesen, J.K., & Allor, J.H. (2001). The effects of Peer Assisted Learning Strategies for first grade readers with and without additional computer assisted instruction in phonological awareness. *American Educational Research Journal, 38,* 371-410.

Mayfield, L.G. (2000). *The effects of structured one-on-one tutoring in sight word recognition of first-grade students at-risk for reading failure*. Unpublished doctoral dissertation, Louisiana Tech University.

McGrady, D. S. (1984). *The effect of programmed tutoring upon the reading comprehension of fourth-grade students enrolled in a Chapter 1reading program.* Paper presented at the

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

annual meeting of the Association for Educational Communications and Technology, Dallas, TX.

Meier, J., & Invernizzi, M. (2001). Book Buddies in the Bronx: Testing a model for America Reads. *Journal of Education for Students Placed at Risk, 6,* 319-333.

Merzenich, M., Jenkins, W., Johnston, P., Schreiner, C., Miller, S., & Tallal, P. (1996). Temporal processing deficits of language learning impaired children ameliorated by training. *Science,* 271, 77-80.

Mooney, P. J. (2003) *An investigation of the effects of a comprehensive reading intervention on the beginning reading skills of first graders at risk for emotional and behavioral disorders*. Ph.D. dissertation, The University of Nebraska – Lincoln.

Morris, D., Tyner, B., & Perney, J. (2000). Early Steps: Replicating the effects of a first-grade reading intervention program. *Journal of Educational Psychology, 92* (4), 681-693.

National Assessment of Educational Progress (2007). *The nation's report card.* Washington, DC: National Center for Education Statistics.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Rockville, MD: National Institute of Child Health and Human Development.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Nelson, J., & Stage, S. (2007). Fostering the development of vocabulary knowledge and reading comprehension through contextually-based multiple meaning vocabulary instruction. *Education and Treatment of Children, 30* (1), 1-22.

Pinnell, G., DeFord, D., & Lyons, C. (1988). *Reading Recovery: Early intervention for at-risk first graders.* Arlington: ERS.

Pinnell, G.S., Lyons, C.A., DeFord, D.E., Bryk, A.S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly 29*(1), 8-39.

Policy Exchange (2009). *Every Child a Reader: An example of how top-down education reforms make matters worse*. Retrieved March 24, 2009, at www.policyexchange.org.uk.

Pullen, P., Lane, H., & Monaghan, M. (2004). Effects of a volunteer tutoring model on the early literacy development of struggling first grade students. *Reading Research and Instruction, 43*(4), 21-40.

Ramey, M. (1991). *Compensatory education sustained gains from spring1988 to spring 1990.* Seattle, WA: Seattle Public Schools.

Ritter, G. (2000). *The academic impact of volunteer tutoring in urban public elementary schools: Results of an experimental design evaluation*. Unpublished doctoral dissertation, University of Pennsylvania.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Ritter, G., Barnett, J., Denny, G., & Albin, G. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research, 79* (1), 3-38.

Ross, S.M., & Casey, J. (1998b). *Success for All evaluation, 1997-98 Tigard-Tualatin School District.* Memphis: University of Memphis, Center for Research on Educational Policy.

Ross, S.M., Nunnery, J.A., & Smith, L.J. (1996). *Evaluation of Title I reading programs: Amphitheater Public Schools Year 1: 1995-1996.* Memphis, TN: University of Memphis, Center for Research in Educational Policy.

Ross, S., Smith, L., Bond, C., Casey, J., & Johnson (1993). *Final report: 1994-94 Success for All program in Montgomery, AL.* Memphis, TN: Center for Research on Educational Policy.

Ross, S.M., Smith, L.J., & Casey, J. (1995). *Final Report: 1994-95 Success for All program in Fort Wayne, Indiana.* Memphis: University of Memphis, Center for Research in Educational Policy.

Ross, S.M., Smith, L.J., & Casey, J. (1994). *Final report: 1993-94 Caldwell, Idaho. SFA results*. Memphis, TN: The University of Memphis. Center for Research in Educational Policy

Rothstein, H.R., Sutton, A.J., & Borenstein, M. (Eds.) (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustments.* Chichester, UK: John Wiley.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Rouse, C.E., & Krueger, A.B. (2004). Putting computerized instruction to the test: A randomized evaluation of a "scientifically-based" reading program. *Economics of Education Review, 23* (4), 323-338.

Santa, C.M. & Høien, T. (1999). An assessment of early steps: A program for early intervention of reading problems. *Reading Research Quarterly, 34*, 54-79.

Schwartz, R.M. (2005). Literacy learning of at-risk first-grade students in the Reading Recovery early intervention. *Journal of Educational Psychology, 97* (2), 257-267.

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin, 105*, 309-316.

Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly, 30*, 958-995.

Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher, 15* (9), 5-11.

Slavin, R. (2008). What works? Issues in synthesizing education program evaluations. *Educational Researcher, 37* (1), 5-14.

Slavin, R.E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best evidence synthesis. *Reading Research Quarterly, 43* (3), 290-322.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Slavin, R.E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (in press). Effective reading programs for the elementary grades. *Review of Educational Research.*

Slavin, R.E., & Madden, N. A. (in press). Bias due to measures inherent to treatments in systematic reviews in education. *Journal of Research on Educational Effectiveness.*

Slavin, R.E., & Madden, N.A. (Eds.) (2009). *Two million children: Success for All*. Thousand Oaks, CA: Corwin.

Slavin, R.E., Madden, N.A., Dolan, L.J., & Wasik, B.A. (1993). *Success for All in the Baltimore City Public Schools: Year 6 report. Baltimore, MD:* Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.

Slavin, R.E., & Smith, D. (in press). Effects of sample size on effect size in systematic reviews in education. *Educational Evaluation and Policy Analysis*.

Smith, L.J., Ross, S.M., & Casey, J. (1994). *Final report: 1993-94 Success For All program in Memphis, Tennessee.* Memphis, TN: University of Memphis, Center for Research on Educational Policy.

Snow, C.E., Burns, S.M., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative Integrated Reading and Composition: Two field experiments. *Reading Research Quarterly, 22*, 433-454.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Stevens, R.J., & Slavin, R.E. (1995a). The cooperative elementary school: Effects on students' achievement, attitudes and social relations. *American Educational Research Journal,* 32(2), 321-351.

Stevens, R., & Slavin, R.E. (1995b). Effects of a cooperative learning approach in reading and writing on academically handicapped and non-handicapped students. *The Elementary School Journal, 95 (*3), 241-261.

Stevens, R., Van Meter, P., Garner, J., Warcholak, N., Bochna, C., & Hall, T. (2008). Reading and Integrated Literacy Strategies (RAILS): An integrated approach to early reading. *Journal of Education for Students Placed at Risk, 13* (4), 357-380.

Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., Nagarajan, S. Shchreiner, C., Jenkins, W., Merzenich, M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, 271, 81-84.

Torgesen, J., Schirm, A., Castner, L., Vartivarian, S., Mansfield, W., Myers, D., Stancavage, F., Durno, D., Javorsky, R., & Haan, C. (2007). *National assessment of Title I final report. Volume II: Closing the reading gap: Findings from a randomized trial of four reading interventions for striving readers.* Washington, DC: IES.

Torgesen, J.K., Wagner, R.K., & Rashotte, C.A. (1997). Prevention and remediation of severe reading disabilities: Keeping the end in mind. *Scientific Studies of Reading, 1,* 217-234.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Torgesen, J.K., Wagner, R.K., Rashotte, C.A., Herron, J., & Lindamood, P. (2009). *Computer assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches*. Tallahassee, FL: Florida State University.

Torgerson, C. J., Brooks, G., & Hall, J. (2006). A systematic review of the research literature on the use of phonics in the teaching of reading and spelling. *London: DfES Research Report 711.*

Vadasy, P.F., Sanders, E.A. (2008). Benefits of repeated reading intervention for low-achieving fourth- and fifth-grade students. *Remedial and Special Education, 29*, 235-249.

Vadasy, P.F., & Sanders, E.A. (2008). Repeated reading intervention: Outcomes and interactions with readers' skills and classroom instruction. *Journal of Educational Psychology, 100* (2), 272-290.

Vadasy, P.F., Sanders, E.A., & Peyton, J. (2005). Relative effectiveness of reading practice or word-level instruction in supplemental tutoring: How text matters. *Journal of Learning Disabilities, 38* (4), 364-380.

Vadasy, P.F., Sanders, E.A., & Tudor, S. (2007). Effectiveness of paraeducator-supplemented individual instruction: Beyond basic decoding skills. *Journal of Learning Disabilities, 40* (6), 508-525.

Vadasy, P.F., Wayne, S., O'Connor, R., Jenkins, J., Pool, K., Firebaugh, M., & Peyton, J. (2005). *Sound Partners: A tutoring program in phonic-based early reading*. Longmont, CO: Sopris West.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Vaughn, S., Linan-Thompson, S., Kouzekanani, K., Bryant, D.P., Dickson, S., & Blozis, S.A. (2003). Reading instruction groups for students with reading difficulties. *Remedial and Special Education, 24*(5), 301-315.

Vernon-Feagans, L., Amendum, S., Kainz, K., Ginsberg, M., Wood, T., & Bock, A. (2009, March). *The Targeted Reading Intervention (TRI): A classroom teacher Tier 2 intervention to help struggling readers in early elementary school*. Paper presented at the annual meetings of the Society for Research on Effective Education, Crystal City, VA.

Wasik, B.A. & Slavin, R.E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28,* 178-200.

Wasik, B.A. (1997). Volunteer tutoring programs: Do we know what works? *Phi Delta Kappan, 79* (4), 282-287.

Webb, N. (2008). Cooperative learning. In T. L. Good (Ed.), *21st Century Education: A reference handbook.* Thousand Oaks, California: Sage.

What Works Clearinghouse (2009). *Beginning reading. What Works Clearinghouse Topic Report.* Retreived March 24, 2009, from http://ies.ed.gov/ncee/wwc/.

Williams, H., Curry, J., & Griffith, J. (1995). *Reading Recovery in AISD*. Austin Independent School District, Department of Audit and Evaluation.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Table 7

Ratings of Strength of Evidence of Individual Programs

Strong Evidence of Effectiveness

Success for All (CIP + TT)

Direct Instruction/Corrective Reading (CIP, SGT)

Peer Assisted Learning Strategies (PALS) (CIP)

Reading Recovery (TT)

Targeted Reading Intervention (TT)

Quick Reads (SGT)

One-to-One Teacher Tutoring with Phonics Emphasis * (TT)
- -Auditory Discrimination in Depth (TT)
- -Early Steps/Howard Street Tutoring (TT)
- -Intensive Reading Remediation (TT)
- -Reading Rescue (TT)
- -Reading with Phonology (TT)

One-to-One Paraprofessional/Volunteer Tutoring with Phonics Emphasis * (T-Para/Volunteer)
- -Sound Partners (T-Para)
- -The Reading Connection (T-Para)
- -SMART (T-Para)
- -Reading Rescue (T-Para)
- -Howard Street Tutoring(T-Para)
- -Book Buddies (T-Volunteer)

* These are similar one-to-one tutoring programs. Each has evidence of effectiveness from studies with sample sizes too small to qualify for "strong evidence" on their own.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Moderate Evidence of Effectiveness

Cooperative Integrated Reading and Composition (CIRC) (CIP)

Limited Evidence of Effectiveness: Strong Evidence of Modest Effects

Jostens (CAI)

Limited Evidence of Effectiveness: Weak Evidence with Notable Effects

Contextually-Based Vocabulary Instruction (CIP)

Early Intervention in Reading (SGT)

Edmark (T-Para)

Empower Reading (SGT)

Lexia (CAI)

Lindamood Phoneme Sequence Program (SGT)

Precision Teaching (CIP)

Proactive Reading (SGT)

Programmed Tutorial Reading (T-Para)

Project READ (CIP)

RAILS (CIP)

Read Naturally (SGT)

Read, Write, and Type (SGT)

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Styles (CIP)

Responsive Reading (SGT)

Same Age Tutoring (CIP)

SHIP (SGT)

TEACH (TT)

Voyager Passport (SGT)

Wallach and Wallach (T-Para)

○ Insufficient Evidence of Effectiveness

Academy of Reading (CAI)

Destination Reading (CAI)

Experience Corps (T-Para)

Failure-Free Reading (SGT)

Fast ForWord (CAI)

Gottshall Small Group Phonics (SGT)

Headsprout (CAI)

HOSTS (T-Volunteers)

New Heights (SGT)

Knowledge Box (CAI)

LeapTrack (CAI)

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Plan Focus (CAI)

Read 180 (CAI)

Spell Read (SGT)

Targeted Intervention (SGT)

Waterford (CAI)

Wilson Reading (SGT)

Key:

TT: One-to-one tutoring by certified teachers

T-Para: One-to-one tutoring by paraprofessionals

T-Volunteer: One-to-one tutoring by volunteers

SGT: Small group tutorials

CIP: Classroom Instructional Process

CAI: Computer-assisted instruction

N No Qualifying Studies

100 Book Challenge

A Comprehensive Curriculum for Early Student Success (ACCESS)

Academic Associates Learning Centers

Accelerated Reader

ALEKS®

ALPHabiTunes

Alpha-Phonics

Balanced Early Literacy Initiative

Barton Reading and Spelling System

Benchmark

BookMARK

Bradley Reading and Language Arts

Breakthrough to Literacy

Bridge

Bridge to Reading

Bring the Classics to Life

CIERA School Change Framework

Comprehensive Early Literacy Learning

Classwide Peer Tutoring©

Compensatory Language Experiences and Reading Program (CLEAR)

Core Knowledge

Cornerstone Literacy Initiative

Curious George Reading and Phonics

DaisyQuest

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Davis Learning Strategies™

Discover Intensive Phonics for Yourself

Discovery World

Dominie

Dr. Cupp Readers® & Journal Writers

Early Success

Early to Read

Earobics®

Emerging Readers

Essential Skills

Evidence Based Literacy Instruction

Exemplary Center for Reading Instruction (ECRI)

Fast Track Action

Felipe's Sound Search

First grade Literacy Intervention Program (FLIP)

First Steps

Flippen Reading Connections™

Fluency Formula

FOCUS: A Reading and Language Program

Four Block Framework

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Frontline Phonics

Fundations

Funnix

GOcabulary Program for Elementary Students

Goldman-Lynch Language Stimulation Program

Goldman-Lynch Sounds-in-Symbols

Great Leaps

Guided Discovery LOGO

Guided Reading

Harcourt Accelerated Reading Instruction

Higher Order Thinking Skills (HOTS)

Hooked on Phonics®

Huntington Phonics

IntelliTools Reading

Insights: Reading as Thinking

Invitations to Literacy

Irlen method

Jigsaw Classroom

Johnny Can Spell

Jolly Phonics

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Kaleidoscope

KidCentered learning

Knowledge Box

Ladders to Literacy

Language for Learning

Language for Thinking

Leap into Phonics

Letter People

Letterland

LinguiSystems

Literacy Collaborative

Literacy First

Little Books

Little Readers

LocuTour

Matchword

Merit Reading Software Program

Multicultural Reading and Thinking Program (McRAT)

My Reading Coach

New Century Integrated Instructional System

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Next Steps

Onward to Excellence

Pacemaker

Pacific Literacy

Pause, Prompt, & Praise©

Peabody Language Development Kits

Performance Learning Systems

Phonemic Awareness in Young Children

Phonics for Reading

Phonics Q

Phono-Graphix

PM Plus Readers

Primary Phonics

Programmed Tutorial Reading

Project Child

Project FAST

Project LISTEN

Project PLUS

Rainbow Reading

Read Well

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Bridge

Reading Explorer's Pathfinders Tutoring Kit

Reading Intervention for Early Success

Reading Rods

Reading Step by Step

Reading Success from the Start

Reading Upgrade

Richards Read Systematic Language Program

Right Start to Reading

Road to the Code

ROAR Reading System

S.P.I.R.E.

SAIL (Second grade Acceleration to Literacy)

Saxon Phonics

Schoolwide Early Language and Literacy (SWELL)

Sing, Spell, Read, and Write (SSRW)

SkillsTutor

Soar to Success

Soliloquy

Sonday System

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Sound Reading

Sounds and Symbols Early Reading Program

Spalding Writing Road to Reading

Starfall

Start Up Kit

Stepping Stones to Literacy

STEPS (Sequential teaching of Explicit Phonics and Spelling)

Stories and More

Story Comprehension to Go

Storyteller Guided Reading

Strategies the Work

Student Team Achievement Divisions (STAD)

Successmaker®

Sullivan Program

Super QAR

Teacher Vision®

Ticket to Read

Touchphonics

Tribes learning Communities®

Verticy Learning

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Voices Reading

VoWac (Vowel Oriented Word Attack Course)

WiggleWorks

Wright Skills

Writing to Read

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Appendix 1

Program	Study	Reason for Exclusion
1:2 tutoring	Lennon, J. E., & Slesinski, C. (1999). Early intervention in reading: Results of a screening and intervention program for kindergarten students. School Psychology Review, 28(3), 353–364.	Duration < 12 weeks
Academy of Reading	Fiedorowicz, C. (1986). Training of component reading skills. *Annals of Dyslexia, 36*, 318-34.	Insufficient sample size
Academy of Reading	Goodloe-Johnson , M., McKinley, N., Rose, J., & Kokkinis, A. (2006). Effectiveness of Academy of Reading in CCSD schools. Charleston, SC: CCSD Department of Statistics and Accountability.	Pretest equivalency not established/ documented
Academy of Reading	Wilkinson, Tammy Bruce (2008). The impact of a computer-based reading intervention program, "Academy of Reading" on the reading achievement of second and third graders. Ph.D. dissertation, Mississippi State University, United States -- Mississippi. Retrieved March 12, 2009, from Dissertations & Theses: Full Text database. (Publication No. AAT 3297470).	Duration < 12 weeks
Accelerated Reader	Knox, M. (1996). An experimental study of the effects of The Accelerated Reader Program and a teacher directed program on reading comprehension and vocabulary of fourth and fifth grade students. Dissertation Abstracts International, 57 (10), 4208A (UMI No. 9710798).	No adequate control group; Pretest scores > 1/2 SD apart; Insufficient sample size
After school tutoring	Dunphy, S.K. (2006). Effect of after-school tutoring programs on students' reading achievement and classroom academic reading performance. Unpublished doctoral dissertation, Union University.	No adequate control group; Outside of age/grade parameters
ALL Program (RR but 3-1 instead of 1-1)	Homan, S. , & Hogarty, K. (2001). A Small Group Model for Early Intervention in Literacy: Group Size and Program Effects. , .	No adequate control group; Pretest equivalency not established/ documented
America Reads (tutoring)	Cook, J. (2001). Every moment counts: Pairing struggling young readers with minimally trained tutors. Unpublished doctoral dissertation, Arizona State University.	Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

America Reads (tutoring)	Fitzgerald, J. (2001). Can Minimally Trained College Student Volunteers Help Young At-risk Children to Read Better? *Reading Research Quarterly, 36*(1), 28-47.	No control group
Audiobooks	Stone-Harris, S. (2008). *The benefit of utilizing audiobooks with students who are struggling readers.* Unpublished doctoral dissertation, Walden University.	Duration < 12 weeks
audiotaped Repeated Readings	Conte, R., & Humphreys, R. (1989). Repeated reading: Using audio-taped material enhances oral reading in children with reading difficulties. *Journal of Communication Disorders, 22*, 65-79.	Inadequate outcome measure
Auditory Discrimination in Depth (ADD) & embedded phonics (EP)	Torgesen, J.K., Alexander, A. W., Wagner, R.K., Rashotte, C.A., Voeller, K., Conway, T. & Rose, E. (2001a). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34*, 33-58.	Pretest scores > 1/2 SD apart; Pretest equivalency not established/documented; Insufficient sample size
Auditory Discrimination in Depth (ADD)/ Lindamood Phonemic Sequencing (LiPS)	Alexander, A. Anderson, H., Helilma, P., Voeller, K., & Torgesen, J. (1991). Phonological awareness training and the remediation of analytic decoding deficits in a group of severe dyslexics. *Annals of Dyslexia, 41*, 193-206.	No control group
Auditory Discrimination in Depth (ADD)/Lindamood Phonemic Sequencing (LiPS)	Conway, T., Heilman, P., Gonzalez-Rothi, L., Alexander, A., Adair, J., Crosson B., & Heilman, K. (1998). Treatment of a case of phonological alexia with agraphia using the Auditory Discrimination in Depth (ADD) program. Journal of the International Neuropsychological Society, 4, 608-620.	Outside of age/grade parameters; Inadequate outcome measure
Auditory Discrimination in Depth (ADD)/Lindamood Phonemic Sequencing (LiPS)	Kennedy, K., & Backman, J. (1993). Effectiveness of the Lindamood Auditory Discrimination in Depth Program with students with learning disabilities. Learning Disabilities Research and Practice, 8 (4), 253-259.	Insufficient sample size
Auditory Discrimination in Depth (ADD)/Lindamood Phonemic Sequencing (LiPS)	Simos, P., Fletcher, J., Bergman, E., Breier, J., Foorman, B., Castillo, E., et al. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58*, 1203-1212.	Insufficient sample size
Book Buddies	Baker, D.J. (1998). The effects of a one-to-one reading tutorial program on the reading achievement of first grade students. Unpublished doctoral dissertation, University of Virginia.	Pretest scores > 1/2 SD apart

Book Buddies	Bromley, K., Winters D., & Schlimmer, K. (1994). Book Buddies: Creating enthusiasm for literacy learning. *The Reading Teacher, 47*, 392-400.	No control group
Book Buddies	Fowler, M. , Thacker-Gwaltney, S. , Invernizzi, M. (2002). A Second Year of One-on-One Tutoring: An Intervention for Second Graders with Reading Difficulties. Ann Arbor, MI: CIERA.	No control group
Book Buddies (tutoring)	Abouzeid, M., & Fowler, M. (1998, December). Book Buddies replication study: Ninety miles from Charlottesville, VA: So what? Paper presented at the annual meeting of the National Reading Conference, Austin, TX.	Pretest equivalency not established/ documented
Book Club, Carbo Reading, Early Reading Intervention, Reading Recovery	McIntyre, E., Petrosko, J., Jones, D., Powell, R. Powers, S., Bright, K., & Newsome, F. (2005). Supplemental instruction in early reading: Does it matter for struggling readers? The Journal of Educational Research, 99(2), 99-107.	Demographic differences > 1/2 SD apart
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Bowling Green City Schools 1999-2000).	No adequate control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Campbell County School District).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Charlotte-Mecklenburg Public School District).	Pretest equivalency not established/ documented
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Copperas Cove Independent School District).	No control group

Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Daviess County School District).	No adequate control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (De Zavala Elementary School Fort Worth Independent School District 1998-99).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Fort Worth Independent School District 1999-2000).	Pretest equivalency not established/ documented
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Fulton County Schools).	Pretest equivalency not established/ documented
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Glynn County Schools).	Outside of age/grade parameters
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Hawaii Department of Education).	No control group

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (I.M. Terrell Elementary School Forth Worth Independent School District).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Jersey City Public Schools).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Johnson County School District).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Lawrence Public Schools 2000-01).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Lebanon Community School Corporation).	Pretest equivalency not established/ documented
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Leon County School District).	Pretest equivalency not established/ documented

Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Logan County School District).	Pretest equivalency not established/ documented
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Massillon City School District).	Pretest equivalency not established/ documented
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (New Haven Public Schools).	Outside of age/grade parameters
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Norfolk Public Schools).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Ohio County School District).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Public School 10 Community School District 15).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Public School 27 Community School District 15).	No control group

Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Public School 57 New York City Public Schools).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Pulaski County Schools).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (San Ysidro School District).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (School District of Palm Beach County).	Pretest equivalency not established/ documented
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Sumpter County School District).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. (Available from Breakthrough to Literacy, 2662 Crosspark Rd., Coralville, IA 52241) (Union County Public Schools 2000-2001).	No control group
Breakthrough to Literacy	Breakthrough to Literacy (2003). Submission to the What Works Clearinghouse, topic 1: Interventions for students with beginning reading difficulties. Retrieved from http://www.breakthroughtoliteracy.com/index.html?PHPSESSID=&page=df_lr_studies_mcneill_1 (Bowling Green City Schools, 2001-2002).	Pretest equivalency not established/ documented

Bridge	Biemiller, A., & Siegel, L. (1997). A longitudinal study of the effects of the Bridge reading program for children at risk for reading failure. *Learning Disability Quarterly, 20*(2), 83-92.	Pretest equivalency not established/ documented
BTL, Early Success, Early Intervention, 4Blocks, SRA Reading Mastery, Together we Can	Rightmyer, E.C., McIntyre, E., & Petrosko, J.M. (2006). Instruction, Development, and Achievement of Struggling Primary Grade Readers. Reading Research and Instruction, 45(3), 209-241.	Pretest equivalency not established/ documented
Building Language for Literacy	Flynn, A. (2006). The modified Building Language for Literacy program: A phenomenological study of attitudes towards literacy and individual progress among preschool students with disabilities. Unpublished master's thesis, Pacific Lutheran University.	Inadequate outcome measure
CAI	Grisham, D. , & Wolsey, T. (2008). The Role of Technology in Supporting Struggling Readers. The Role of Technology in Supporting Struggling Readers, 93-115.	Outside of age/grade parameters
CAI	Lin, A., Podell, D., & Rein, N. (1991). The effects of CAI on word recognition in mildly mentally handicapped and nonhandicapped learners. *Journal of Special Education Technology, 11*, 16-25.	Insufficient sample size
CAI	Bentivolio, K. (2001). Improving a student's reading comprehension skills by teaching computer aided design. Educational Resources Information Centre (ERIC) document number ED455507.	Outside of age/grade parameters
CAI	Farmer, M., Klein, R., & Bryson, S. (1992). Computer-assisted reading: effects of whole-word feedback on fluency and comprehension in readers with severe disabilities. *Remedial and Special Education, 13*, 50-60.	Insufficient sample size
CAI	Golden, N., Gersten, R., & Woodward, J. (1990). Effectiveness of guided practice during remedial reading instruction: an application of compuer-managed instruction. *Elementary School Journal, 90*, 291-304.	Outside of age/grade parameters
CAI	Huang, Cheng-Fang (2004) Scaffolding sight vocabulary acquisition for children with autism using computer-assisted instruction. Ed.D. dissertation, University of Washington, United States -- Washington. Retrieved September 4, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3131166).	Inadequate outcome measure

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

CAI	Moseley, D. (1993). Visual and linguistic determinants of reading fluency in dyslexics: A classroom study with talking computers. In S.F. Wright and R. Groner (Eds.), *Facets of dyslexia and its remediation*. Elsevier Science.	No control group
CAI	Olofsson, A. (1992). Synthetic speech and computer aided reading for reading disabled children. *Reading and Writing: An Interdisciplinary Journal, 4,* 165-178.	No control group; Pretest equivalency not established/ documented
CAI	Smeets & Van der Leij (1993) Differential effects of computer assisted instruction in the reading ability of poor and dyslexic readers: A pilot study	Pretest equivalency not established/ documented
CAI	Sutherland, M., & Smith, C. (1997). The benefits and difficulties of using portable word prcoessors with older dyslexics. Dyslexia, 3, 15-26.	Inadequate outcome measure
CAI	Tjus, T., Heimann, M., & Nelson, K. (1998). Gains in literacy through the use of a specially developed multimedia computer strategy. Autism, 2, 139-156.	Inadequate outcome measure
CAI	Yap & Van der Leij (1993) Computer based remeidation of reading disability by sub lexical speed training	Pretest scores > 1/2 SD apart; Pretest equivalency not established/ documented
CAI	Greenleaf, C. (1994). Technological indeterminacy: the role of classroom writing practices and pedagogy in shaping student use of the computer. *Written Communication, 11*, 85-130.	Pretest equivalency not established/ documented
CAI	Jiminez, J., del Rosario Ortiz, M., Rodrigo, M., Hernandez-Valle, I., Ramirez, G., Estevez, A., O'Shanahan, I., & de la Luz Trabaue, M. (2003). Do the effects of computer-assisted practice differ for children with reading disabilities with and without IQ-achievement discrepancy? *Journal of Learning Disabilities, 36*(1), 34-47.	Insufficient sample size
Carbo Reading Styles Program	Bradsby, S., Wise, J., Mundell, S. & Haas, S. (1992). "Making a difference for L. D. students – Matching reading instruction to reading styles through recorded books." Research in the Classroom (ERIC Document Reproduction Service No. ED 347 765).	No control group

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Carbo Reading Styles Program	Mohrmann, S. (1990, January). Learning styles of poor readers. Paper presented at the meeting of the Southwest Educational Research Association, Austin, TX.	Pretest equivalency not established/ documented
Carbo Reading Styles Program	Wilson, I. (1993). Reading styles of Hispanic students with learning disabilities in third, fourth, and fifth grade. Dissertation Abstracts International, 55 (11), 3462A (UMI No. 9505375).	No control group
CATTS (computer-assisted teacher training system)	Semmel, M., Indiana Univ., B., & Others, A. (1976, August 1). The Effectiveness of a Computer-Assisted Teacher Training System (CATTS) in the Development of Reading and Listening Comprehension Instructional Strategies of Preservice Special Education Trainees in a Tutorial Classroom Setting. Final Report 53.4. (ERIC Document Reproduction Service No. ED162467)	Insufficient sample size
CD ROM books	Greenlee-Moore, M. E., & Smith, L. L. (1996). Interactive computer software: The effects on young children's reading achievement. *Reading Psychology, 17,* 43-64.	Insufficient sample size
chapter 1 services	Hiebert, E. H., Colt, J. M., Catto, S. L., & Gury, E. C. (1992). Reading and writing of first grade students in a restructured Chapter I program. *American Educational Research Journal, 29*(3), 545-572.	Pretest scores > 1/2 SD apart
CHAR-L Intensive phonics program	Cobb, S. (1990). Effectiveness of Phonics for an Intensive Remedial Program. Reading Improvement, 27(3), 218-219.	Duration < 12 weeks
ClassWide Peer Tutoring	Perdomo-Rivera, C. (2002). The effects of classwide peer tutoring on the literacy achievement and language production of English Language Learners in an elementary school setting. Unpublished doctoral dissertation, University of Kansas.	No control group
Classwide Peer Tutoring (CWPT)	Burks, M. (2004). Effects of Classwide Peer Tutoring on the number of words spelled correctly by students with LD. *Intervention in School and Clinic, 39*(5), 301-304.	No control group
Classwide Peer Tutoring (CWPT)	Ezell, H.K., Kohler, F.W., & Strain, P.S. (1994). A program description and evaluation of academic peer tutoring for reading skills of children with special needs. Education and Treatment of Children, 15, 205-227.	No control group

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Classwide Peer Tutoring (CWPT)	Sidiridis,G., Utley, C., Greenwood, C., & Delquadri, J. et al. (1997). Classwide Peer Tutoring: Effects of the spelling performance and social interactions of students with mild disabilities and their typical peers in an integrated instructional setting. *Journal of Behavioral Education, 7*(4), 203-212.	No adequate control group
Classwide Peer Tutoring (CWPT)	Greenwood, C., Dinwiddie, G., Bailey, V., Carta, J., Dorsey, D., Kohler, F., Nelson, C., Rotholtz, D., & Schulte, D. (1987). Field replication of classwide peer tutoring. *Journal of Applied Behavior Analysis, 20*, 151-160.	Outside of age/grade parameters
Classwide Peer Tutoring (CWPT)	Buzhardt, J., Greenwood, C., Abbott, M., & Tapia, Y. (2006). Research on scaling up effective instructional intervention practice: Developing a measure of the rate of implementation. *Educational Technology Research and Development, 54*(5), 467-492.	No control group
Classwide Peer Tutoring (CWPT) + LMS	Abbott, M., Greenwood, C.R., Buzhardt, J., & Tapia, Y. (2006). Using technology-based teacher support tools to scale up the ClassWide Peer Tutoring program. *Reading and Writing Quarterly, 22*, 47-64.	No control group
Classwide Peer Tutoring (CWPT) Learning Management System	Buzhardt, J., Abbott, M., Greenwood, C.R., & Tapia, Y. (2005). Usability testing of the ClassWide Peer Tutoring-learning Management System. *Journal of Special Education Technology, 20*(1), 19-31.	Demographic differences > 1/2 SD apart
classwide peer-mediated reading instruction (says like CWPT, never mentions PALS but clearly like it)	Mathes, P.G., & Fuchs, L. (1993). Peer-Mediated Reading Instruction in Special Education Resource Rooms. *Learning Disabilities Research and Practice, (8)*4, 233-43.	No control group
CLEAR-Reading Recovery	Bermel, S. (1987). Language development component, CLEAR-Reading Recovery Program 1985-86. Final evaluation report. Columbus, OH: Columbus Public Schools, Ohio Department of Evaluation Services (ERIC No ED281157).	Pretest scores > 1/2 SD apart
CLIP (or some type of Reading Recoveryesque program)	Alegria-Romero, M.L. (2006). Development and assessment of an early literacy intervention program in an elementary school. Unpublished doctoral dissertation, Northern Arizona University	Pretest scores > 1/2 SD apart

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Cogent-Prep	Hayward, D., Das, J., & Janzen, T. (2007, September). Innovative programs for improvement in reading through cognitive enhancement: a remediation study of Canadian First Nations children. Journal of Learning Disabilities, 40(5), 443-457	Insufficient sample size
Collections/ Intervention Readers	Educational Research Institute of America (ERIA). Winter/Spring 2001 study of the instructional effectiveness of the intervention readers in Harcourt's Reading/Language Arts program: COLLECTIONS c2001	No adequate control group
Compass Learning	Interactive, Inc. (2004). *Pocatello School District: An analysis of CompassLearning student achievement outcomes in Pocatello, Idaho. Research Report.* San Diego, CA: Compass Learning.	No adequate control group
comprehension strategy instruction	Manset-Williamson, G. & Nelson, J.M. (2005). Balanced, strategic reading instruction for upper-elementary and middle school students with reading disabilities: A comparative study of two approaches. *Learning Disability Quarterly, 28,* 59-74.	No control group
cori	Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., Scafiddi, N. T., & Tonks, S. (2004). Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. Journal of Educational Psychology, 96, 403 423.	No control group
corrective reading	Flores, M., & Ganz, J. (2007, Winter2007). Effectiveness of Direct Instruction for Teaching Statement Inference, Use of Facts, and Analogies to Students With Developmental Disabilities and Reading Delays. Focus on Autism & Other Developmental Disabilities, 22(4), 244-251.	No control group
Corrective Reading	Benner, G.J., Kinder, D., Beaudoin, K.M., Stein, M., & Hirschmann, K. (2005). The effects of the *Corrective Reading* Decoding program on the basic reading skills and social adjustment of students with high incidence disabilities. *Journal of Direct Instruction, 5*(1), 67-80.	No control group
cross-age tutoring	Carberry, David John (2003) *The effects of cross-age tutoring in reading on tutees, tutors and metacognitively trained tutors.* Unpublished doctoral dissertation, University of Minnesota.	Insufficient sample size

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

cross-age tutoring	Coats, L.B. (2007). Cross-age tutoring: Effects on reading achievement of tutors and tutees in an after-school program. Unpublished doctoral dissertation, Northcentral University.	Inadequate outcome measure
cross-age tutoring	Standley, L. (2006). *Cross-age peer-tutoring effects on the English literacy development and academic motivation of English language learners identified with, and referred for, mild and moderate disabilities.* Unpublished doctoral dissertation, The University of New Mexico	No control group
CWPT	Bradley, D., Bjorlykke, L., Mann, E., Homon, C., & Lindsay, J. (1993, October). *Empowerment of the general educator through effective teaching strategies.* Paper presented at the meeting of the International Conference on Learning Disabilities, Baltimore, MD.	No adequate control group
CWPT	Neddenriep, Christine Elizabeth (2003). Classwide peer tutoring: Three experiments investigating the generalized effects of increased oral reading fluency to silent reading comprehension. Ph.D. dissertation, The University of Tennessee, United States -- Tennessee. Retrieved March 12, 2009, from Dissertations & Theses: Full Text database. (Publication No. AAT 3104401).	Outside of age/grade parameters
CWPT	Simmons, D., Fuchs, D., Fuchs, L.S., Pate, J., & Mathes, P. (1994). Importance of instructional complexity and role reciprocity to classwide peer tutoring. *Learning Disabilities Research and Practice, 9(4),* 203-212.	Demographic differences > 1/2 SD apart
DaisyQuest	Barker, T., & Torgesen, J. K. (1995). An evaluation of computer-assisted instruction in phonological awareness with below average readers. *Journal of Educational Computing Research, 13*(1), 89–103.	Insufficient sample size
DaisyQuest	Barker, T.A. (1993). An evaluation of computer-assisted instruction in phonological awareness with below-average readers. Unpublished doctoral dissertation, The Florida State University.	Insufficient sample size
DaisyQuest	Mitchell, M. J., & Fox, B. J. (2001). The effects of computer software for developing phonological awareness in low-progress readers. *Reading Research and Instruction, 40*(4), 315–332.	No adequate control group

DARA	Macrine, S., & Sabbatino, E. (2008, January). Dynamic assessment and remediation approach: Using the DARA approach to assist struggling readers. Reading & Writing Quarterly: Overcoming Learning Difficulties, 24(1), 52-76.	Pretest equivalency not established/ documented
Decoding	Pullen, P., Lane, H., & Lloyd, J. (2005). Effects of Explicit Instruction on Decoding of Struggling First Grade Students: A Data-Based Case Study. Education and Treatment of Children, 28(1), 63-75. Retrieved January 27, 2009, from Education Full Text database.	Insufficient sample size
DECtalk	Leong, C. (1995). Effects of on-line reading and simultaneous DECtalk auding in helping below-average and poor readers comprehend and summarize text. *Learning Disability Quarterly, 18*, 101-116.	Inadequate outcome measure
Designed by researchers	VAUGHN, S. , WANZEK, J. , MURRAY, C. , SCAM MACCA, N. , LINAN-THOMPSON, S. , et al. (2009). Response to Early Reading Intervention: Examining Higher and Lower Responders. Exceptional Children, 75(2), 165-183.	Insufficient sample size
DI/Corrective Reading	Lloyd, J., Cullinan, D., Heins, E.D., & Epstein, M. (1980). Direct Instruction: Effects on oral and written language comprehension. Learning Disabilities Quarterly, 3.	Inadequate outcome measure
Direct Instruction	Branwhite, A. (1983, January 1). Boosting Reading Skills by Direct Instruction. British Journal of Educational Psychology, 53(3), 291-98.	Insufficient sample size
Direct Instruction / Corrective Reading	Gregory, R. P., Hackney, C., & Gregory, N. M. (1982). Corrective Reading programme: An evaluation. *British Journal of Educational Psychology, 52,* 33–50.	Inadequate outcome measure
Direct Instruction/ Corrective Reading	Polloway, E., Epstein, M., Polloway, C., Patton, J., & Ball,D. (1986). Corrective Reading program: An analysis of effectiveness with learning disabled and mentally retarded students. *Remedial and Special Education, 7*(4), 41-47.	No control group
Direct Instruction/ Corrective Reading	Campbell, M. (1984). Corrective Reading program evaluated with secondary students in San Diego. ADI News, 3, 3.	Outside of age/grade parameters

Direct Instruction/ Corrective Reading	Harris, R., Marchand-Martella, N., Martella, R. (2000). Effects of a peer-delivered Corrective Reading program. *Journal of Behavioral Education, 10*, 21-36.	No control group
Direct Instruction/ Corrective Reading	Slaton, D. (2006). Effects of Corrective Reading on the reading abilities and classroom behaviors of middle school students with reading deficits and challenging behavior. *Behavioral Disorders, 313*, 265-283.	Outside of age/grade parameters
Direct Instruction/ Corrective Reading	Somerville, D., & Leach, D. (1988, February). Direct or indirect instruction: An evaluation of three types of intervention programs for assisting students with specific reading difficulties. Educational Research, 30 (1), 46-53.	No adequate control group
Direct Instruction/ Corrective Reading	Vitale, M., Medland, M., Romance, N., & Weaver, H. P. (1993). Accelerating reading and thinking skills of low-achieving elementary students: Implications for curricular change. *Effective School Practices, 12*(1), 2-31.	Pretest equivalency not established/ documented
Direct Instruction/SRA	Koehler, K. (1996). The effects of phonological awareness and letter naming fluency on reading acquisition for first-graders experiencing difficulty learning to read. Dissertation Abstracts International, 57 (7), 2944A (UMI No. 9638095).	Insufficient sample size
Discrete vs continuous speech	Higgins, E., & Raskind, M. (2000). Speaking to read: the effects of continuous vs discrete speech recognition systems on the reading and spelling of children with learning disabilties. *Journal of Special Education Technology, 15*(1), 19-30.	Insufficient sample size
DISTAR	Kuder, S. (1991). Language Abilities and Progress in a Direct Instruction Reading Program for chidrenwith learning disabilities" Journal of learning disabilities, 24 (2)124 - 127.	No control group
Distar and Johnny Right-to-Read	Summerell, S., & Brannigan, G.G. (1977). *Comparison of reading programs for children with low levels of reading readiness. Perceptual and Motor Skills, 44*(3), 743-6.	Inadequate outcome measure
DISTAR vs Integrated Reading-Writing	Traweek, D., & Berninger, V. (1997). Comparisons of beginning literacy programs: Alternative paths to the same learning outcome. *Learning Disability Quarterly, 20*(2), 160-168.	No control group

DISTAR/Corrective Reading vs Integrative Skills method	Richardson, E., Dibenedetto, B., Christ, A., Press, M., & Winsbert, B. (1978). An assessment of two methods for remediating reading deficiencies. *Reading Improvement, 15*(2), 82-95.	no control group
Early identification	Hurford, D.P., Johnston, M., Nepote, P., Hampton, S., Moore, S., Neal, J., Mueller, A., McGeorge, K., Huff, L., Awad, A., Tatro, C., Juliano, C., & Huffman, D. (1994). Early identification and remediation of phonological-processing deficits in first-grade children at risk for reading disabilities. *Journal of Learning Disabilities, 27*(10), 647-659.	Insufficient information on outcome data
Early Intervention in Reading (EIR)	Wing, M.A. (1994). The Effects of a Supplemental Literacy Program on Students in a Developmental First-Grade Classroom Using Cross-age Tutors. Dissertation Abstracts International, 50 (1), 151A (UMI No.9514687).	Inadequate outcome measure
Early Intervention in Reading (EIR) / Coss-Age tutoring	Taylor, B.M., Hanson, B., Justice-Swanson, K. & Watts, S. (1997). Helping Struggling Readers: Linking Small-Group Intervention with Cross-Age Tutoring. *The Reading Teacher, 51*(3), 196-209.	Inadequate outcome measure
Early Literacy Project (Project Read a large component)	Englert, C., Garmon, A., Mariage, T., Rozendal, M., Tarrant, K., & Urba, J. (1995). The Early Literacy Project: Connecting across the literacy curriculum. *Learning Disability Quarterly, 18*(4), 253-275.	Pretest scores > 1/2 SD apart
earobics	Gale, D. (2006). *The effect of computer-delivered phonological awareness training on the early literacy skills of students identified as at-risk for reading failure.* Retrieved from the University of South Florida website: http://purl.fcla.edu/usf/dc/et/SFE0001531.	No control group
Earobics	Pobanz, M. (2000, January). The effectiveness of an early literacy/auditory processing training program, called Earobics, with young children achieving poorly in reading. Paper presented at the meeting of the California Association of Social Psychologists, Los Angeles, CA.	No control group
Earobics	Pobanz, M. (2003). Outcomes report: Los Angeles Unified School District. Evanston, IL: Cognitive Concepts.	No control group
earobics	Rehmann, R. (2005). *The effect of Earobics (TM) Step 1, software on student acquisition of phonological awareness skills.* Unpublished doctoral dissertation, Univeristy of Oregon.	Pretest equivalency not established/ documented

earobics	Valliath, S. (2002). An evaluation of a computer-based phonological awareness training program: Effects on phonological awareness, reading and spelling. Unpublished doctoral dissertation, Northwestern University.	Duration < 12 weeks
EIP - early intervention program	McCarthy, P., Newby, R. F., & Recht, D. R. (1995). An early intervention program for first grade children at-risk for reading disability. *Reading research and Instruction, 34*, 273-294.	Pretest equivalency not established/ documented
elaborative interrogation	Abdulaziz, T. M. (1999). The role of elaborative interrogation in acquiring knowledge from expository prose passages for students with learning and behavior disorders. *Dissertation Abstracts International, 61*(01A), 132.	Demographic differences > 1/2 SD apart
ElectroText	Horney, M., Anderson Inman, L. (1999). Supported text in electronic reading environments. *Reading and Writing Quarterly: Overcoming Learning Diffiiculties, 15*, 127-168.	Inadequate outcome measure
Elements of Reading: Comprehension	Resendez, M., Sridiharan, S., & Azin, M. (2006). Harcourt Achieve's Elements of Reading: Comprehension radomized control trial. PRES associates.	No adequate control group
elements of reading:phonics and phonemic awareness	Apthorp, H. (2005). Elements of Reading: Phonics and Phonemic Awareness. Orlando: Harcourt.	Pretest equivalency not established/ documented
ENABLE-Plus	Bowen, P., & Yeomans, J. (2002). Accelerating the Progress of Failing Readers: An Evaluation of the ENABLE-Plus Programme Pilot Study. *British Journal of Special Education, 29* (4), 170-177. EJ659326	No control group
Epi-Meta-Mastery Approach	Frost, Y. & Sørensen, P.M. (2007). The effects of a comprehensive reading intervention programme for Grade 3 children. *Journal of Research in Reading 30*(3), 270–286.	Pretest equivalency not established/ documented
ERI	Singh, Angella Harjani (2008). Follow-up study of the effects of a supplemental early reading intervention on the reading skills of urban at-risk primary learners. Ph.D. dissertation, The Ohio State University	No control group
Essential Learning Systems	Holmes, S. (2001). *The Relative Effectiveness of Essential Learning Systems, a Sensory Integration Training Program on Introductory Reading Skills and Academic Self-Concept of Rural African American Children with Learning Deficits.* Unpublished doctoral dissertation, University of Mississippi.	Insufficient sample size

explicit instruction vs basal instruction	Rabren, K., Darch, C., & Eaves, R.C. (1999). The differential effects of two systematic reading comprehension approaches with students with learning disabilities. *Journal of Learning Disabilities, 32,* 36-47.	No control group
explicit instruction…peer tutoring part of interest (PALS)	Simmons, D., Fuchs, L., & Fuchs, D., Mathes, P., & Hodge, J.P. (1995). Effects of explicit teaching and peer tutoring on the reading achievement of learning disabled and low-performing students in regular classrooms. *Elementary School Journal, 95*(5)*,* 387-408.	Pretest equivalency not established/ documented
Failure Free Reading	Lockavitch, J.F., & Algozzine, B. (1998). Effects of intensive intervention on students at risk for reading failure. *The Florida Reading Quarterly, 35*(2), 27-31.	Pretest equivalency not established/ documented
Failure Free Reading	Algozzine, B., & Lockavitch, J. (1998). Effects of the Failure Free Reading program on students at risk for reading failure. *Special Services in the Schools, 13* (1/2), 95-103.	No control group
Failure Free Reading	Algozzine, B., Lockavitch, J., & Audette, R. (1997). Implementing Failure-Free Reading with students seriously at-risk for failure. Australian Journal of Learning Disabilities, 2 (3), 14-17.	No control group
Failure Free Reading	Bergquist, C., Richardson, G., Bigbie, C., Castine, W., Hancock, W., Largent, W. et al. (2001). Final report of the Failure Free Reading Bridges programs funded under Florida's 2000 Specific Appropriation 5A: Executive summary. Tallahassee, FL: Evaluation Systems Design, Inc.	No adequate control group
Failure Free Reading	Blount, L. (2003). Clay County School District comprehensive school reform grant project summary and evaluation report July 1, 1998-June 30, 2001. Green Cove Springs, FL: Clay County School District.	No adequate control group
Failure Free Reading	Educational Enhancement Services. (2000). Greensboro Elementary School comprehensive school reform evaluation report. Retrieved August 26, 2006, from http://www.failurefree.com/downloads/Greensboro_CSRD_Report.pdf.	No adequate control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Chester Elementary).	No control group

Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Fullerton Elementary).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Lincoln Elementary).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Lowest literacy students during OhioReads).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Lyme Elementary).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: North Elementary, Urbana City Schools).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Perry Elementary).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: SC Dennis Elementary).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Williamson Elementary).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. Retrieved August 26, 2006 from http://www.failurefree.com/downloads/FFR_OHReads_Set_1.pdf. (Study: Midway Elementary).	No control group

Failure Free Reading	Failure Free Reading (2003). Failure Free reading research findings: OhioReads 2000-01 school year results. Retrieved August 26, 2006 from http://www.failurefree.com/downloads/FFR_OH Reads_Set_1.pdf. (Study: Miles Standish Elementary).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free Reading's continuum of effectiveness: Research summary (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025)(Study: Greensboro Elementary, Gadsden County, FL).	No control group
Failure Free Reading	Failure Free Reading (2003). Failure Free Reading's continuum of effectiveness: Research summary (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025) (Study: Helen S. Edwards Elementary, New Orleans, LA)	No control group
Failure Free Reading	Failure Free Reading (2003). OhioReads research evaluation (2000-2001 School Year) impact on lowest literacy students. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Lowest literacy students during OhioReads).	No control group
Failure Free Reading	Failure Free Reading (n.d.). Case study: Fairland East Elementary's after-school solution. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025).	No control group
Failure Free Reading	Failure Free Reading (n.d.). Program effectiveness has been shown through an experimental design that includes experimental and control groups created through random assignment or carefully matched comparison groups. Retrieved from http://www.failurefree.com/downloads/FFR_vs_ Control.pdf. (Study: Copperas Cove, ISD)	Pretest equivalency not established/ documented
Failure Free Reading	Failure Free Reading (n.d.). Program effectiveness has been shown through an experimental design that includes experimental and control groups created through random assignment or carefully matched comparison groups. Retrieved from http://www.failurefree.com/downloads/FFR_vs_ Control.pdf. (Study: Cowee Elementary, Macon County, NC).	No adequate control group

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Failure Free Reading	Failure Free Reading (n.d.). Program effectiveness has been shown through an experimental design that includes experimental and control groups created through random assignment or carefully matched comparison groups. Retrieved from http://www.failurefree.com/downloads/FFR_vs_Control.pdf. (Study: Southwest Elementary).	Pretest equivalency not established/ documented
Failure Free Reading	Failure Free Reading (n.d.). Program effectiveness has been shown through an experimental design that includes experimental and control groups created through random assignment or carefully matched comparison groups. Retrieved from http://www.failurefree.com/downloads/FFR_vs_Control.pdf. (Study: Washington, DC - River Terrace & Miner Elementary)	Pretest equivalency not established/ documented
Failure Free Reading	Failure Free Reading (n.d.). Research summary intensive intervention for upper elementary students. Retrieved from http://www.failurefree.com/downloads/FFR_Upper_Elem_Intervention.pdf (Study: Florida CSRD Sites).	No control group
Failure Free Reading	Failure Free Reading. (1999). Failure Free Reading's Impact on North Carolina's end of grade assessment. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025) (Study: Catawba County).	No adequate control group
Failure Free Reading	Failure Free Reading. (1999). Failure Free Reading's Impact on North Carolina's end of grade assessment. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025) (Study: Johnson County-Benson Elementary, NC).	No adequate control group
Failure Free Reading	Failure Free Reading. (1999). Failure Free Reading's Impact on North Carolina's end of grade assessment. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025) (Study: Lincoln County).	No adequate control group
Failure Free Reading	Failure Free Reading. (1999). Failure Free Reading's Impact on North Carolina's end of grade assessment. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025) (Study: Rutherford County-Futherfordton Elementary, NC).	No adequate control group

Failure Free Reading	Failure Free Reading. (1999). Four week summer school with Failure Free Reading produces greater growth than entire year. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025).	No adequate control group
Failure Free Reading	Failure Free Reading. (1999). Twelve days with Failure Free Reading produced dramatic results. (North Carolina Research Brief 99.102). Retrieved from http://www.failurefree.com/downloads/FFR_Catawba.pdf	Outside of age/grade parameters
Failure Free Reading	Failure Free Reading. (2003). Case study: Fairland East Elementary's fourth grade reading blitz. Concord, NC: Author. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025).	No control group
Failure Free Reading	Failure free Reading. (2003). Failure Free Reading research findings: Intervention for beginning reading. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Greenwood, MS)	No control group
Failure Free Reading	Failure free Reading. (2003). Failure Free Reading research findings: Intervention for beginning reading. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Rowan County, NC: Reading readiness study of at-risk first graders).	No control group
Failure Free Reading	Failure Free Reading. (2003). Failure Free Reading's continuum of effectiveness: Research summary. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Fairland East Elementary, Proctorville, OH).	Duration < 12 weeks
Failure Free Reading	Failure Free Reading. (2003). Failure Free Reading's continuum of effectiveness: Research summary. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: West Clay Elementary, Clay County, MS).	No control group

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Failure Free Reading	Failure Free Reading. (2003). Washington, DC--Reed Elementary 2002/03 results. Retrieved from http://www.failurefree.com/downloads/FFR_Reed_Elem_2003.pdf.	No control group
Failure Free Reading	Failure Free Reading. (2004). Supplemental educational service provider (SSP): Bacon School, Millville, NJ. Retrieved from http://www.failurefree.com/downloads/Bacon_Results_Summary.pdf	No control group
Failure Free Reading	Failure Free Reading. (n.d.). Dramatic intensive intervention results in Chicago. Retrieved from: http://failurefree.com/downloads/Dulles_Elem_Chicago.pdf	No control group
Failure Free Reading	Failure free Reading. (n.d.). Failure Free Reading research findings: OhioReads 2000-01 school year results. Retrieved August 26, 2006 from http://www.failurefree.com/downloads/FFR_OHReads_Set_1.pdf (Study: Hamden Elementary).	No control group
Failure Free Reading	Failure free Reading. (n.d.). Failure Free Reading research findings: OhioReads 2000-01 school year results. Retrieved August 26, 2006 from http://www.failurefree.com/downloads/FFR_OHReads_Set_1.pdf (Study: Seacrest Elementary).	No control group
Failure Free Reading	Failure free Reading. (n.d.). Failure Free Reading research findings: OhioReads 2000-01 school year results. Retrieved August 26, 2006 from http://www.failurefree.com/downloads/FFR_OHReads_Set_1.pdf (Study: Shumaker Elementary).	No control group
Failure Free Reading	Failure Free Reading. (n.d.). Research findings concerning the impact of the Failure Free Reading program on at-risk and special education lowest literacy students. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Accelerated growth curve).	No control group

Failure Free Reading	Failure Free Reading. (n.d.). Research findings concerning the impact of the Failure Free Reading program on at-risk and special education lowest literacy students. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Learning curve of at-risk and special education students).	No control group
Failure Free Reading	Failure Free Reading. (n.d.). Research findings concerning the impact of the Failure Free Reading program on at-risk and special education lowest literacy students. (Available from Failure Free Reading, 140 Cabarrus Ave., W., Concord, NC 28025). (Study: Sustaining growth).	No control group
Failure Free Reading	Failure Free Reading. (n.d.). Research summary intensive intervention for upper elementary students. Retrieved from http://www.failurefree.com/downloads/FFR_Upper_Elem_Intervention.pdf. (Study: Klein ISD).	No control group
Failure Free Reading	Failure Free Reading. (n.d.). Research summary intensive intervention for upper elementary students. Retrieved from http://www.failurefree.com/downloads/FFR_Upper_Elem_Intervention.pdf. (Study: Washington, DC-Spring 2002).	Insufficient sample size
Fast ForWord	Borman, G. D., & Benson, J. (2006). Can brain research and computers improve literacy? A randomized field trial of the Fast ForWord Language computer-based training program (WCER Working Paper No. 2006-5). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research.	Insufficient sample size
Fast ForWord	Breier, J., Gray, L., Fletcher, J., Diehl, R., Klass, P., Foorman, B., et al. (2001). Perception of voice and tone onset time continua in children with dyslexia with and without attention deficit/hyperactivity disorder. *Journal of Experimental Child Psychology, 80*(3), 245-270.	Demographic differences > 1/2 SD apart
Fast ForWord	Habib, M., Espesser, R., Rey, V., Giraud, K., Brunas, P., & Gres, C. (1999). Training dyslexics with acoustically modified speech: Evidence of improved phonological awareness. *Brain & Cognition, 40*, 143-146.	No control group

Fast ForWord	Hook, P., Macaruso, P., & Jones, S. (2001). Efficacy of Fast ForWord training on facilitating acquisition of reading skills by children with reading difficulties: A longitudinal study. *Annals of Dyslexia, 51*, 75-96.	Insufficient sample size
Fast ForWord	Marler, J., Champlin, C., & Gillam, R. (2001). Backward and simultaneous masking measured in children with language-learning impairments who received intervention with Fast ForWord or Laureate Learning Systems Software. *American Journal of Speech Language Pathology, 10*(3), 258-268.	No control group
Fast ForWord	Merzenich, M., Jenkins, W., Johnston, P., Schreiner, C., Miller, S., & Tallal, P. (1996). Temporal processing deficits of language learning impaired children ameliorated by training. *Science, 271*, 77-80.	Demographic differences > 1/2 SD apart
Fast ForWord	Merzenich, M., Miller, S., Jenkins, W., Saunders, G., Protopapas, A., Peterson, B., & Tallal, P. (1997). Amelioration of the acoustic reception and speech reception deficits underlying language based learning impairments. In C. von Euler, I. Lundberg, & R. Linas (Eds.), Basich mechanisms in cognition and language (pp. 143-172). New York: Elsevier.	No control group
Fast ForWord	Merzenich, M., Tallal, P., Peterson, B., Miller, S., & Jenkins, W. (1999). Some neurological principles relevant to the origins of--and the cortical plasticity-based remediation of--developmental language impairments. In J. Grafman & Y. Christen (Eds.), Neuroplasticity: Building a bridge from the laboratory to the clinic. (pp. 169-187). Amsterdam: Elsevier.	Pretest scores > 1/2 SD apart
Fast ForWord	Schopmeyer, B., Mellon, N., Dobaj, H., Grant, G., & Niparko, J. (2000). Use of Fast ForWord to enhance language development in children with cochlear implants. Annals of Otology, Rhinology, & Laryngology, 109 (12), 95-98.	No control group
Fast ForWord	Scientific Learning Corporation. (2004). Improved language and reading skills by students in the School District of Philadelphia who were receiving services for special education and who used Fast ForWord products. Maps to Learning: Educator Reports, 8 (20), 1-4.	Pretest scores > 1/2 SD apart

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Fast ForWord	Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., Nagarajan, S. Shchreiner, C., Jenkins, W., Merzenich, M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271*, 81-84.	No control group
Fast ForWord	Troia, G., & Whitney, S. (2002). A close look at the efficacy of Fast ForWord Language for children with academic weaknesses. *Contemporary Educational Psychology, 28*(4), 465-494.	Insufficient sample size; Demographic differences > 1/2 SD apart
FastStart Reading	Rasinski, T. & Stevenson, B. (2005). The effects of Fast Start Reading: A fluency-based home involvement reading program on the reading achievement of beginning readers. Reading Psychology: An International Quarterly, 26 (2), 109-125.	Inadequate outcome measure
Fluency Formula	Sivin-Kachala, J., & Bialo, E. (2005). Fluency Formula second grade study. Long Island, NY, New York: scholastic.	Inadequate outcome measure
four blocks	Scarcelli, S., & Morgan, R. (1999). The efficacy of using a direct reading instruction approach in literature based classrooms. Reading Improvement, 36 (4), 172-179.	Pretest equivalency not established/ documented
general CAI	Calvert, S., Watson, J., Brinkley, V., & Penny, J. (1990). Computer presentational featuers for poor readers' recall of infromation. Journal of Educational Computing Research, 6, 287-298.	No control group; Insufficient sample size
Graphic Organizers	DiCecco, V., & Gleason, M. (2002). Using graphic organizers to attain relational knowledge from expository texts. Journal of Learning Disabilities, 35, 306-320.	Outside of age/grade parameters
Guided Reading	Cornejo, Jessica Renee (2007) Guided reading: An effective intervention for struggling readers. M.S. dissertation, California State University	Duration < 12 weeks; Insufficient sample size
harcourt vs di	Clark, D. E., Jr. (2007) A comparison of 3rd grade reading scores between students using an integrated literature based curriculum and students using direct instruction at a charter school. Unpublished doctoral dissertation, Wilmington College.	No control group; Pretest equivalency not established/ documented
Headsprout	Clarfield, J., & Stoner, G. (2005). The effects of computerized reading instruction on the academic performance of students identified with ADHD. *School Psychology Review, 34*, 246-255.	No control group

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Hint and Hunt	Jones, K., Torgesen, J., & Sexton, M. (1987). Using computer guided practice to increase decoding fluency in learning disabled children: a study using the Hint and Hunt 1 program. *Journal of Learning Disabilities, 20*, 122-128.	No adequate control group; Pretest scores > 1/2 SD apart; Insufficient sample size
Hoffman Language Arts, Ginn Reading 720	Carlton, S.B. (1981). *Reading achievement, student attitude, and program costs:a comparative study of two programed supplementary reading programs.* Unpublished doctoral dissertation, The Florida State University.	Inadequate outcome measure
Horizon Fast Track/ PALS	Barton-Arwood, S.M. (2003). *Reading instruction for elementary-age students with emotional and behavioral disorders: Academic and behavioral outcomes.* Unpublished doctoral dissertation, Vanderbilt University.	No control group
HOSTS	Bradley, K.L. (2001). *The effects of the Help One Student to Succeed (HOSTS) program on the reading achievement of at-risk 4th and 5th grade elementary students.* Unpublished doctoral dissertation, Old Dominion University.	No adequate control group; Pretest scores > 1/2 SD apart
Howard Street Tutoring Program	Morris, D., Shaw, B., & Perney, J. (1990). Helping low readers in grades 2 and 3: An after-school volunteer tutoring program. *The Elementary School Journal, 91*(2), 133-191.	No control group; Insufficient sample size; Demographic differences > 1/2 SD apart
ICARE (tut)	Hedrick, D.E. (1996). *An administrative review of an early reading intervention.* Unpublished doctoral dissertation, University of North Carolina at Greensboro.	No control group
K2	Robinson-Evans, J.M. (2006). *An investigation of the effects of an early reading intervention on students with disabilities and those at-risk of reading failure*. Unpublished doctoral dissertation, Ball State University.	Pretest equivalency not established/ documented
K-PALS (not really a study of this though…brief experimental analyses (BEAs))	Petursdottir, A.G. (2006) Brief experimental analysis of early reading interventions. Unpublished doctoral dissertation, University of Minnesota.	Pretest equivalency not established/ documented; Insufficient sample size
Ladders to Literacy	Notari-Syverson, A., O'Connor, R., & Vadasy, P. (1996). Supporting the development of early literacy in preschool children with disabilities. Seattle: Washington Research Institute.	No control group
Ladders to Literacy	O'Connor, R. (1999). Teachers Learning Ladders to Literacy. Learning Disabilities Research & Practice, 14(4), 203-214.	Insufficient sample size

Ladders to Literacy	O'Connor, R., Notari-Syverson, A., & Vadasy, P. (1996, March). The effect of kindergarten phonological intervention on the first grade reading and writing of children with mild disabilities. Paper presented at the annual meeting of the American Educational Research Association, New York.	Pretest equivalency not established/ documented; Insufficient sample size
Lighthouse (after school)	Farmer-Hinton, R. (2002). When Time Matters: Examining the Impact and Distribution of Extra Instructional Time. Proceedings of the Annual Meeting of the National Association of African American Studies, National Association of Hispanic and Latino Studies, National Association of Native American Studies, and International Association of Asian Studies, Houston, TX, Feb. 11-16, 2002.	Pretest equivalency not established/ documented
Lightspan/Achieve Now	Gwaltney, L. (2000). Year three final report the Lightspan Partnership, Inc. Achieve Now Project: Unified School District 259, Wichita Public Schools. Wichita, KS: Allied Educational Research and Development Services.	Treatment confounded with other programs
Lindamood-Bell method	Tergeson, J.K, Wagner, R.K., Rashotte, C.A., Alexander, A.W., & Conway, T. (1997). Preventive and remedial interventions for children with severe reading disabilities. *Learning Disabilities: A multidisciplinary Journal, 8.*	No control group
literacy coaching	Schuster, R. (2004). Professional development and student literacy: A program evaluation of literacy coaching. Unpublished doctoral dissertation, Saint Louis University.	Outside of age/grade parameters
metacognitive strategy training	Greaney, K., Tunmer, W., & Chapman, J. (1997). Effects of rime-based orthographic analogy training on the word recognition skills of children with reading disability. Journal of Educational Psychology, 89, 645-651.	No adequate control group
methods for teaching spelling…not reading	Vaughn, S., Schumm, J., & Gordon, J. (1993). Which motoric condition is most effective for teaching spelling to students with and without learning disabilities? Journal of Learning Disabilties, 26, 191-198.	No control group; Pretest equivalency not established/ documented
modified whole language	Eldredge, L. (1991). An experiment with a modified whole language approach in first-grade classrooms. *Reading Research and Instruction, 30*, 21-38.	Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Montessori	Ibeji, N.O. (2002) Improving early reading skills of first-grade students with learning disabilities using Montessori learning strategies. Ph.D. dissertation, Union Institute and University	Insufficient sample size
MultiFunk	Fasting, R. B. & Lyster, S. H. (2005). The effects of computer technology in assisting the development of literacy in young struggling readers and spellers. *European Journal of Special Needs Education, 20*(1), 21–40.	Duration < 12 weeks
Multiple Connections Model	Berninger, V.	No adequate control group
My Reading Coach	Bliss, J., Larrabee, J., & Schnitzler, P. (2002). The performance of a new computer-based reading tutor. Retrieved from Mindplay web site: http://images/pcmac.org/Uploads/ELSSystems/ELSSytstems/Divisons/DocumentsCategories/Documents/Comp-basedReadingTeacher.pdf	No control group
one-to-one tutoring (pa)	Gibbs, S. E. L. (2001). *Effects of a one-to-one phonological awareness intervention on first grade students identified as at risk for the acquisition of beginning reading.* Unpublished doctoral dissertation, University of South Carolina.	No control group; Duration < 12 weeks
Onset Rime decoding	Hines, S. (2009). The Effectiveness of a Color-Coded, Onset-Rime Decoding Intervention with First-Grade Students at Serious Risk for Reading Disabilities. Learning Disabilities Research & Practice (Blackwell Publishing Limited), 24(1), 21-32.	Insufficient sample size
PALS	Calhoon, M., Otaiba, S., Greenberg, D., King, A., & Avalos, A (2006). Improving reading skills in predominately Hispanic Title I first grade classrooms: The promise of Peer-Assisted Learning Strategies. Learning Disabilities Research and Practice, 21 (4), 261-272.	Pretest equivalency not established/ documented
PALS	Lane, K.L., Wehby, J., Menzies, H.M., Gregg, R.M., Doukas, G.L., Munton, S.M. (2002). Early Literacy Instruction for First-Grade Students At-Risk for Antisocial Behavior Education & Treatment of Children, Vol. 25, 2002	Outside of age/grade parameters; Inadequate outcome measure
PALS	Lorah, K.S. (2003). *Effects of peer tutoring on the reading performance and classroom behavior of students with attention deficit hyperactivity disorder.* Unpublished doctoral dissertation, Lehigh University.	Insufficient sample size; Inadequate outcome measure

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

PALS	Mathes, P.G., Grek, M.L., Jill K. Howard, Allison E. Babyak, , Shelley H. Allen. (1999) Peer-Assisted Learning Strategies for First-Grade Readers: A Tool for Preventing Early Reading Failure. Learning Disabilities Research and Practice 14:1, 50-60	No control group
Parental involvement	Ellis, M. (1996). The Effects of a Parent-child Reading Program on Reading Ability and Self-perceptions of Reading Ability in Struggling Young Readers. Dissertation Abstracts International Section A: Humanities and Social Sciences, 57(3), 1013.	No control group
Partners in Reading and Reading Recovery	Miller, S. (2003). Partners-in-Reading: Using classroom assistants to provide tutorial assistance to struggling first-grade readers. *Journal of Education for Students Placed at Risk, 8*(3), 333-349.	Demographic differences > 1/2 SD apart
PATR	Lane,K.L., O'Shaughnessy, T.E., Lambros,K.M., Gresham, F.M. & Beebe-Frankenberger, M.E. (2001). The efficacy of phonological awareness training with first-grade students who have behavior problems and reading difficulties . Journal of Emotional and Behavioral Disorders, 9(4), 219-232.	No control group
PAYC, Reading Readiness, Ladders to Literacy, PATR	Wanzek, J., Shirley Dickson, William D. Bursuck , Jennifer M. White. (2000) Teaching Phonological Awareness to Students At Risk for Reading Failure: An Analysis of Four Instructional Programs. Learning Disabilities Research and Practice 15:4, 226-239	No control group; Demographic differences > 1/2 SD apart
peer tutoring	August, D.L. (1987). Effects of peer tutoring on the second language acquisition of Mexican American children in elementary school. *TESOL Quarterly, 21*(4), 717-736.	Inadequate outcome measure
Peer tutoring	Pruitt, R. (2008). The Effects of a Literacy-based Service-learning Program on Struggling Fourth Grade Readers. Dissertation Abstracts International Section A: Humanities and Social Sciences, 68(9), 3740.	Pretest equivalency not established/ documented; Insufficient sample size
Peer-Assisted Learning Strategies (PALS)	Allor, J., Fuchs, D., & Mathes, P. (2001). Do students with and without lexical retrieval weaknesses respond differently to instruction? *Journal of Learning Disabilities, 34*(3), 264-275.	Pretest scores > 1/2 SD apart

Peer-Assisted Learning Strategies (PALS)	Baker, R. (2005). Teacher Directed Instruction Plus Classwide Peer Tutoring and the Reading Growth of First Grade Students. Unpublished masters thesis, California State University, Fresno.	Pretest scores > 1/2 SD apart
Peer-Assisted Learning Strategies (PALS)	Fuchs, L., Fuchs, D., Kazdan,S., Allen, S. (1999). Effects of Peer-Assisted Learning Strategies in Reading with and Without Training in Elaborated Help Giving. Elementary School Journal, 99(3), 201-220.	Insufficient sample size
Peer-Assisted Learning Strategies (PALS)	Hudson, K.G. (2004). The effects of Peer-Assisted Learning Strategies on the reading achievement of elementary students with and without decoding weaknesses. Unpublished doctoral dissertation, University of Virginia.	Demographic differences > 1/2 SD apart
Peer-Assisted Learning Strategies (PALS)	Pearson, J.J.M. (2004). The effect of peer-assisted literacy strategies on the social standing of first-grade readers. Unpublished doctoral dissertation, University of Houston.	Demographic differences > 1/2 SD apart
Peer-Assisted Learning Strategies (PALS)	Saenz, L., Fuchs, L., & Fuchs, D. (2005). Peer-Assisted Learning Strategies for English language learners with learning disabilities. *Exceptional Children, 71*, 231-247.	Demographic differences > 1/2 SD apart
PHAB/DI and WIST	Lovett, M. W. & Steinbach, K. A. (1997). The effectiveness of remedial program for reading disabled children of different ages: Does the benefit decrease for older children? *Learning Disability Quarterly, 20*(3), 189-210.	No control group; Insufficient sample size
PHAB/DI and WIST	Lovett, M.W., Borden, S.L., DeLuca, T., Lacerenza, L., Benson, N.J., & Brackstone, D. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategically-based training programs. *Developmental Psychology, 30*(6), 805-822.	No control group; Insufficient sample size
phoneme awareness training	Ball, E.W. & Blachman, B.A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and development spelling? *Reading Research Quarterly, 26*(1), 49-66.	Insufficient sample size
phoneme training, onset rime training, whole word training	Haskell, D., Foorman, B., & Swank, P. (1992). Effects of three orthographic/phono- logical units on first-grade reading. Remedial and Special Education, 13, 40-49.	Insufficient sample size

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Phonemic awareness and phonemically based decoding skills	Ryder, J. , & Greaney, K. (2008). Explicit Instruction in Phonemic Awareness and Phonemically Based Decoding Skills as an Intervention Strategy for Struggling Readers in Whole Language Classrooms. Reading and Writing: An Interdisciplinary Journal, 21(4), 349-369.	Insufficient sample size
phonemic awareness training	Bentin, S. & Leshem, H. (1993). On the interaction between phonological awareness and reading acquisition: It's a two-way street. *Annals of Dyslexia, 43,* 125-148.	Demographic differences > 1/2 SD apart
phonemic awareness training	Blachman, B. A., Tangel, D. M., Ball, E. W., Black, R., & McGraw, C. K. (1999). Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. Reading and Writing: An Interdisciplinary Journal, 11, 239-273.	Insufficient sample size
phonemic awareness training	Ryder, J., Tunmer, W., & Greaney, K. (2008, June). Explicit instruction in phonemic awareness and phonemically based decoding skills as an intervention strategy for struggling readers in whole language classrooms. Reading and Writing, 21(4), 349-369.	No control group; Pretest equivalency not established/ documented; Insufficient sample size
phonemic awareness training	Blachman, B. A., Ball, E., Black, R., & Tangel, D. (1994). Kindergarten teachers develop phoneme awareness in low-income inner-city classrooms: Does it make a difference? Reading and Writing: An interdisciplinary Journal 6: 1-17.	Pretest scores > 1/2 SD apart; Pretest equivalency not established/ documented
Phonics for Reading	Boone, B.A. (2004). A reading intervention for first grade students at-risk for reading failure. Unpublished doctoral dissertation, California State University, Fresno.	Insufficient sample size
Phono-Graphix	McGuiness, C., McGuinness, C., & McGuinness, G. (1996). Phono-Graphix: A new method for remediating reading difficulties. Annals of Dyslexia, 46, 73-96.	Pretest scores > 1/2 SD apart
Phono-Graphix	Endress, S. A., Weston, H., Marchand-Martella, N.E., Martella, R.C., & Simmons, J. (2007). Examining the effects of Phono-Graphix on the remediation of reading skills of students with disabilities: a program evaluation. *Education & Treatment of Children, 30*(2), 1-20.	No control group; Insufficient sample size

Phono-Graphix and Read Naturally	Denton, C., Fletcher, J., Anthony, J., & Francis, D. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities, 39*(5), 447-466.	No adequate control group; Insufficient sample size; Demographic differences > 1/2 SD apart
Phonological Awareness	Pokorni, J. , Worthington, C. , & Jamison, P. (2004). Phonological Awareness Intervention Comparison of Fast ForWord, Earobics, and LiPS. The Journal of Educational Research (Washington, D.C.) V. 97 No. 3 (January/February 2004) P. 147-57, 97(3), 147-157.	No control group; Duration < 12 weeks
phonological awareness training	Lie, A. (1991). Effects of a training program for stimulating skills in word analysis in first-grade children. *Reading Research Quarterly, 26*, 234-250.	No adequate control group; Pretest equivalency not established/ documented
Phonological Awareness Training	Blumsack, J. (1996). Teaching phonological awareness to children with language impairments. Unpublished doctoral dissertation, Syracuse University.	Insufficient sample size
Phonological Awareness Training for Reading	Asfendis, George (2008) Phonemic awareness and early intervention: An evaluation of a pilot phonemic awareness program. Psy.D. dissertation, Fairleigh Dickinson University, United States -- New Jersey. Retrieved January 12, 2009, from Dissertations & Theses: Full Text database. (Publication No. AAT 3327689).	Pretest scores > 1/2 SD apart
Phonological Decoding	Wentink, H. Van Bon, W., & Schreuder, R. (1997). Training of Poor Readers' Phonological Decoding Skills: Evidence for Syllable-Bound Processing. Reading and Writing: An Interdisciplinary Journal, 9, 163-192.	Insufficient sample size
phonological training - no specific program though	Torgesen, J.K., Morgan, S., & Davis, C. (1992). Effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology, 84,* 364-370.	Insufficient sample size
Phonological training: SP/LPA vs. LPA	Del Rosario Ortiz Gonzalez, M., Garcia Espinel, A. I., & Guzman Rosquete, R. (2002). Remedial interventions for children with reading disabilities: Speech perception—An effective component in phonological training? *Journal of Learning Disabilities, 35*(4), 334-342.	Insufficient sample size

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

phrase-cued text	Johnson, J.L. (2007). The Use of Phrase-cued Text as an Intervention to Facilitate Oral Reading Fluency for Struggling Third Graders. Unpublished doctoral dissertation, University of South Dakota.	Insufficient sample size
PLATO/WICAT	Erdner, R., Guy, R., & Bush, A. (1997). The impact of a year of computer assisted instruction on the development of first grade reading skills. Journal of Educational Computing Research, 18 (4), 369-388.	Insufficient sample size
Precision Reading	Freeze. R. (2004). Precision Reading with reluctant readers in a French Immersion elementary school. Unpublished research report. Available from the author, Faculty of Education, University of Manitoba.	No control group; Pretest equivalency not established/ documented
Precision Reading	Updike, M, & Freeze, R. (2002). Precision Reading: Improving reading for students with leraning disabilities.	Pretest equivalency not established/ documented
Programmed Reading	Kamps, D. , Abbott, M. , Greenwood, C. , Wills, H. , Veerkamp, M. , et al. (2008). Effects of Small-Group Reading Instruction and Curriculum Differences for Students Most at Risk in Kindergarten. Journal of Learning Disabilities, 41(2), 101-114.	Insufficient sample size
Project LISTEN's Reading Tutor	Aist, G. (2001). Towards automatic glossarization: Automatically constructing and administering vocabulary assistance factoids and multiple-choice assessment. International Journal of Artificial Intelligence in Education, 12, 212-231.	No adequate control group
Project LISTEN's Reading Tutor	Aist, G., & Mostow, J. (2000, June). Using automated within-subject invisible experiments to test the effectiveness of automated vocabulary assistance. Paper presented at the meeting of the Workshop on Modeling Human Teaching Tactics and Strategies, Montreal, Canada.	No adequate control group
Project LISTEN's Reading Tutor	Aist, G., Kort, B., Reilly, R., Mostow, J., & Picard, R. (2002, June). Experimentally augmenting an intelligent tutoring system with human supplied capabilities: Adding human-provided emotional scaffolding to an automated reading tutor that listens. Paper presented at the meeting of the Workshop on Empirical Methods for Tutorial Dialogue Systems, San Sebastian, Spain.	No adequate control group; Inadequate outcome measure

Project LISTEN's Reading Tutor	Beck, J., Jia, P., & Mostow, J. (2003, June). Assessing student proficiency in a reading tutor that listens. Paper presented at the meeting of the International Conference on User Modeling, Johnstown, PA.	No control group
Project LISTEN's Reading Tutor	Beck, J., Mostow, J., Cuneo, A., & Bey, J. (2003, July). Can automated questioning help children's reading comprehension? Paper presented at the meeting of the International Conference on Artificial Intelligence in Education, Sydney, Australia.	No control group
Project LISTEN's Writing Tutor	Mostow, J., Beck, J., Bey, J., Cuneo, A., Sison, J., Tobin, B., et al. (2004). Using automated questions to assess reading comprehension, vocabulary, and effects of tutorial intervention. *Technology, Instruction, Cognition and Learning, 2*, 103-140.	Pretest equivalency not established/ documented
Project LITT	Lewis, R. (2000). Project LITT enhancing the reading skills of students with learning disabilities through hypermedia-based chidlren's literature: Final report. San Diego, CA: San Diego State University, CA, Dept of Special Education.	No adequate control group; Pretest equivalency not established/ documented; Inadequate outcome measure
Project PRIDE	Bursuck, W., Smith, T., Munk, D., Damer, M., Mehlig, L., & Perry, J. (2004). Evaluating the Impact of a Prevention-Based Model of Reading on Children Who Are At Risk. *Remedial and Special Education, 25* (5), 303-313.	Pretest equivalency not established/documented
Project Read	Enfield, M. (1976). *An alternate classroom approach to meeting special learning needs of children with reading problems.* Unpublished masters thesis, University of Minnesota.	No control group
Project Read (comprehension streategies portion)	Cox, D.J. (1997). *The effectiveness of Project Read and visualization and verbalization reading comprehension strategies to improve reading comprehension in at-risk and learning disabled students.* Unpublished master's thesis, California State University, Fresno.	Pretest scores > 1/2 SD apart
Questioning the Author	Beck, I., McKeown, M., Worthy, J., Sandora, C., & Kucan, L. (1996). Questioning the author: A year-long classroom implementation to engage students with text. Elementary School Journal, 96, 385-414.	No control group
R.A.T. Pack	Butler, S.R. (1991). Reading program - remedial, integrated, and innovative. Annals of dyslexia, 41, 119-127.	Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Rainbow Reading Program	Nalder, S., (2002). The effectiveness of Rainbow Reading: An audio-assisted reading program. Retrieved from Pacific Learning web site: http://www.pacificlearning.com/Pages/articles/NHRReffectiveness1.doc.	No control group
read naturally	Jitendra, A. K., Edwards, L. L., Starosta, K., Sacks, G., Jacobson, L. A., & Choutka, C. (2004). Early reading instruction for children with reading difficulties: Meeting the needs of diverse learners. Journal of Learning Disabilities, 27, 421–439.	Pretest scores > 1/2 SD apart; Pretest equivalency not established/ documented
Read Naturally	Read Naturally. (n.d.). Case 1: Original study, Minneapolis, MN. Retrieved from http://www.readnaturally.com/approach/case1.htm.	No control group; Pretest equivalency not established/ documented; Inadequate outcome measure
Read Naturally	Read Naturally. (n.d.). Case 2: Special education students. Huron County, MI. Retrieved April 25, 2007, from http://www.readnaturally.com/why/case2.htm	No control group; Pretest equivalency not established/ documented
Read Naturally	Read Naturally. (n.d.). Case 9: Special education students. Upper Lake, CA. Retrieved April 25, 2007 from http://www.read-naturally.com/why/case9.htm.	No adequate control group
Read Naturally (TMRR - teacher modeling/repeated reading strategy)	Ihnot, C., & Marston, D. (1990). Using teacher modeling and repeated reading to improve the reading performance of mildly handicapped students. Unpublished master's thesis, Minneapolis, University of Minnesota.	No adequate control group; Pretest scores > 1/2 SD apart
Read Well	Santoro, L. , Starosta, K. , & Sacks, G. (2006). Reading Well with "Read Well": Enhancing the Reading Performance of English Language Learners. Remedial and Special Education, 27(2), 105-115.	Insufficient sample size
Read Well/ Read Naturally	Denton, C. A., Anthony, J. L., Parker, R., & Hasbrouck, J. E. (2004). Effects of two tutoring programs on the English reading development of Spanish-English bilingual students. *The Elementary School Journal, 104*(4), 289–305.	Insufficient sample size

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Readers Theater	Sullivan, C.J. (2007). *Reading to students, script-writing and readers' theatre: Strategies to enhance reading skills of low-achievers in a third-grade early intervention classroom.* Unpublished doctoral dissertation, Capella University.	No control group
reader's theater	Corcoran, C. A. (2005). A study of the effects of readers' theater on second and third grade special education students' fluency growth. Reading Improvement 42(2), 105-111.	No adequate control group; Pretest scores > 1/2 SD apart
Readers' Theater	Mountford, K.A. (2007). *Increase reading fluency of 4th and 5th grade students with learning disabilities using Readers' Theater.*	No control group
Reading Acceleration Program (RAP)	Feazell, V.S. (2004). Reading Acceleration Program: A schoolwide intervention. The Reading Teacher, 58(1), 66-72.	Pretest equivalency not established/ documented
Reading CAT (tutoring with computer)	Chambers, B., Abrami, P., McWhaw, L, & Therrien, M. (2001). Developing a computer assisted tutoring program to help children at risk learn to read. *Educational Research and Evaluation, 7*, 223-239.	Demographic differences > 1/2 SD apart
reading machine	Abram, S.L. (1984). *The effect of computer assisted instruction on first grade phonics and mathematics achievement computation.* Unpublished doctoral dissertation, Northern Arizona University.	Insufficient sample size
Reading Mastery Fast Cycle & Horizons Fast Track A-B	Cooke, N. L., Gibbs, S. L., Campbell, M. L., & Shalvis, S. L. (2005). A comparison of Reading Mastery Fast Cycle and Horizons Fast Track A-B on the reading achievement of students with mild disabilities. Journal of Direct Instruction, 4(2), 139-151.	No control group; Demographic differences > 1/2 SD apart
Reading Recovery	Ansay, P. (1999). Reading Performances of Former Reading Recovery Students. Unpublished doctoral dissertation, University of Massachusetts, Lowell.	Pretest equivalency not established/ documented
Reading Recovery	Ashdown, J. & Simic, O. (2003). Is early literacy intervention effective for English language learners? Evidence from Reading Recovery. In S. Forbes & C. Briggs (Eds.), Research in Reading Recovery (pp. 18-38). Portsmouth, NH: Heinemann.	Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Recovery	Ashdown, J., & Simic, O. (2000). Is early literacy intervention effective for English Language Learners? Evidence from Reading Recovery. *Literacy Teaching and Learning: An International Journal of Early Reading and Writing, 5,* 27–42.	Pretest equivalency not established/ documented
Reading Recovery	Askew, B., & Frasier, D. (1997). Sustained effects of Reading Recovery intervention on the cognitive behaviors of second grade children and the perceptions of their teachers. In. S.L. Swartz & A.F. Klein (Eds.), Research in Reading Recovery (pp. 18-38). Portsmouth, NH: Heinemann.	Pretest equivalency not established/ documented
Reading Recovery	Baenen, N., Bernholc, A., Dulaney, C. & Banks, K. (1997). Reading Recovery: Long-Term Progress after Three Cohorts. *Journal of Education for Students Placed at risk, 2*(2), 161-181.	Inadequate outcome measure
Reading Recovery	Batten, P. (2004, Winter). Investing equity funding in early literacy. *ERS Spectrum, 22*(1), 40–45.	No control group
Reading Recovery	Begoray, D. (2001). The Literacy Groups Project: Investigating the Use of Reading Recovery Techniques with Small Groups of Grade 2 Students. Alberta Journal of Educational Research, 47(2), 141-155.	Insufficient sample size
Reading Recovery	Briggs, C. & Young, B.K. (2003). Does Reading Recovery work in Kansas? A retrospective longitudinal study of sustained effects. *Journal of reading Recovery, 3*(1), 59-64.	No control group; Pretest equivalency not established/ documented
Reading Recovery	Brown, K.L. (1999). The impact of Reading Recovery intervention on the reading achievement of selected second grade students. *Dissertation Abstracts International, 62* (08), 2636A. (UMI No. 3023650).	No control group; Pretest equivalency not established/ documented
Reading Recovery	Brown, T. (2007). *The lasting effects of the Reading Recovery Program on the reading achievement on at risk youth.* Unpublished doctoral dissertation, Capella University.	Pretest equivalency not established/ documented
Reading Recovery	Brown, W., Denton, E., Kelly, P.R., & Neal J.C. (1999). Reading Recovery effectiveness: A five-year success story in San Luis Costal Unified School District. ERS Spectrum 17, 3-12.	Pretest equivalency not established/ documented
Reading Recovery	Bufalino, J. (1993). The sustained effects of Reading Recovery intervention on the reading comprehension of second graders. Dissertation Abstracts international, 54 (11), 145A. (UMI No. 9407866).	Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Recovery	Bufalino, J. Wang, C., & Gomez-Bellenge, F. (2007). What's Possible for First Grade At-Risk Literacy Learners Receiving Early Intervention Services. Paper presented at the 2007 meeting of the American Educational Research Association, Chicago, IL.	Pretest equivalency not established/ documented
Reading Recovery	Busbee, N.W. (2001).	No control group
Reading Recovery	Caraway, M.A.H. (2006). A cross-sectional study of performance on high-stakes state assessment by at-risk students who were served in an early intervention program. Unpublished doctoral dissertation, Texas Woman's University.	No adequate control group
Reading Recovery	Center for Early Literacy	No control group
Reading Recovery	Chapman, J., Tunmer, W., & Prochnow, J. (2001). Does success in the Reading Recovery program depend on developing proficiency in phonological processing skills? A longitudinal study in a whole language instructional context. *Scientific Studies of Reading, 5*(2), 141-176.	Pretest equivalency not established/ documented
Reading Recovery	Christman, M. (2003). *An examination of the effects and costs of the Reading Recovery Program in an urban school district.* Unpublished doctoral dissertation, University of Rochester.	No control group; Pretest equivalency not established/ documented
Reading Recovery	Collins, E. (2000). *The immediate and sustained effects of the Reading Recovery program on grade one and grade four at-risk students: A longitudinal study.* Unpublished doctoral dissertation, Univeristy of Illinois at Urana-Champaign.	Pretest equivalency not established/ documented
Reading Recovery	Collins, V. (1994). *Automaticity in information processing.* Unpublished doctoral dissertation, Georgia State University.	Pretest equivalency not established/ documented
Reading Recovery	Concha, J. S. (2005). *Reading Recovery children and early literacy development: Investigation into phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing.* University of Maryland, College Park: Department of Curriculum and Instruction.	No adequate control group
Reading Recovery	Curtin, J. (1993). The effectiveness of Reading Recovery Program on reading achievement. Chicago: Chicago Public Schools. (ERIC No. ED363863).	No control group; Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Recovery	DeFord, D. (1997). Early writing: Teachers and children in Reading Recovery. In S. Swartz & A. Klein (Eds.), Research in Reading Recovery (pp. 148-172). Portsmouth, NH: Heinemann.	No adequate control group
Reading Recovery	Denton, C. (1997). An Evaluation of an Implementation of the Reading Recovery Program. Paper presented at the Annual Meeting of the Southwest Educational Research Association, Austin, TX. Jan. 23-25, 1997.	No adequate control group; Inadequate outcome measure
Reading Recovery	Department of Defense Education Activity (1998).	Pretest equivalency not established/ documented
Reading Recovery	Department of Evaluation Services. (1995). Compensatory Education (CE) product evaluation: Elementary and secondary programs 1994-95. Saginaw, MI: Saginaw Public Schools. (ERIC No. ED391853).	No adequate control group
Reading Recovery	Dorn, L., & Allen, A. (1995). Helping low-achieving first-grade readers: A program combining Reading Recovery tutoring and small-group instruction. *Journal of Reading Recovery, 13*(3), 16–24.	No control group; Pretest equivalency not established/ documented
Reading Recovery	Douëtil, J. (2004). The long term effects of Reading Recovery on national curriculum tests at end of key stages 1 and 2. London: Institute of Education.	No control group
Reading Recovery	Evans, T.L.P. (1996). *I can read deze books!: A qualitative comparison of the Reading Recovery program and a small group reading intervention.* Unpublished doctoral dissertation, Auburn University.	No control group; Inadequate outcome measure
Reading Recovery	Fitzgerald, J. & Ramsbotham, A. (2004). First graders' cognitive and strategic development in Reading Recovery reading and writing. *Reading Research and Instruction (44)*1, 1-31.	Inadequate outcome measure
Reading Recovery	Gilmer, Vicki Bryan (2003) Sustained success of former Reading RecoveryRTM students. Ph.D. dissertation, Auburn University, United States -- Alabama. Retrieved September 5, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3081577).	Pretest equivalency not established/ documented
Reading Recovery	Gómez-Bellengé, F. & Rodgers, E. (2003). Reading Recovery and Descubriendo la Lectura national report 2001-2002 . Columbus: The Ohio State University, National Data Evaluation Center.	No adequate control group; Pretest scores > 1/2 SD apart

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Recovery	Gómez-Bellengé, F. & Rodgers, E. (2004). Reading Recovery and Descubriendo la Lectura national report 2002-2003 . Columbus: The Ohio State University, National Data Evaluation Center.	No adequate control group; Pretest scores > 1/2 SD apart
Reading Recovery	Hovest, C.M. (2000). *An examination of the achievement of phonological skills for three groups participating in an early intervention program.* Unpublished doctoral dissertation, The Ohio State University,	Pretest equivalency not established/ documented
Reading Recovery	Iversen, S. , Tunmer, W. , & Chapman, J. (2005). The Effects of Varying Group Size on the Reading Recovery Approach to Preventive Early Intervention. Journal of Learning Disabilities, 38(5), 456-472.	No adequate control group
Reading Recovery	Jamison, G.R. (2008). *A longitudinal study of the sustained gains of former discontinued Reading Recovery students.* Unpublished doctoral dissertation, University of Arkansas at Little Rock, Arkansas.	Pretest equivalency not established/ documented
Reading Recovery	LaFave, C.E. (1995). *Impact of Reading Recovery on phonemic awareness.* Unpublished doctoral dissertation, The University of Toledo.	Pretest scores > 1/2 SD apart; Inadequate outcome measure
Reading Recovery	Murphy, J.A. (2003). *An application of growth curve analysis: The evaluation of a reading intervention program.* Unpublished doctoral dissertation, Northern Illinois University.	No control group; Demographic differences > 1/2 SD apart
reading recovery	O'Connor, E., & Simic, O. (2002). The effect of Reading Recovery on special education referrals and placements. *Psychology in the Schools, 39*(6), 635-646.	No control group; Demographic differences > 1/2 SD apart
Reading Recovery	Pinnell, G. S. (1989). Reading Recovery: Helping at-risk children learn to read. The Elementary School Journal, 90(2), 161–183.	Pretest scores > 1/2 SD apart; Pretest equivalency not established/ documented; Inadequate outcome measure
reading recovery	Pinnell, G. S. (1988). Success of children at risk in a program that combines writing and reading. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED 292 061)	No control group; Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Recovery	Pollock, J. (1993). Final evaluation report: Reading Recovery program 1991-92. Columbus, OH: Columbus Public Schools, Ohio Department of Program Evaluation. (ERIC No. ED358440).	No control group
Reading Recovery	Pollock, J. (1996). Reading Recovery Program. Final Evaluation Report 1995-96. Columbus, OH: Columbus Public Schools, Ohio Department of Program Evaluation. ,	No control group
Reading Recovery	Pollock, J. (with Morgan, K., Williams, E., & Amorose, R) (1991). Reading Recovery program 1990-91. Final evaluation report. Elementary and Secondary Education Act--Chapter 1, Columbus, OH: Columbus Public Schools, Ohio Department of Education (ERIC No. ED343108).	No control group
Reading Recovery	Pollock, J., Williams, E., Morgan, K., & Amorose, R. (1990). Language development component compensatory language experiences and reading. CLEAR-Reading Recovery program, 1989-90. Final Evaluation Report. Columbus, OH: Columbus Public Schools (ERIC No. ED327821).	Outside of age/grade parameters
Reading Recovery	Pollock, J.S., Williams, E.J., Morgan, K.L., Amorose, R.A. (1993). Final evaluation report, Reading Recovery program, 1991-1992. Columbus (Ohio) Public Schools, Department of Program Evaluation.	Pretest scores > 1/2 SD apart
Reading Recovery	Quay, L.C., Steele, D., Johnson, C., & Hortman, W. (2001). Children's achievement and personal and social development in a first-year Reading Recovery program with teachers in training. *Literacy Teaching and Learning, 5*, 7-25.	Pretest equivalency not established/ documented
Reading Recovery	Reading Recovery Program 1996-97 Evaluation Report. Dept. of Defense Education Activity, Arlington, VA.	Pretest scores > 1/2 SD apart; No adequate control group
Reading Recovery	Rodgers, E., Gomez-Bellenge, F., & Fullerton, S. (2003). Reading Recovery in Ohio: 2001-2002 state report. (National Data Evaluation Center Tech. Rep. No. 2003-03). Columbus, OH: Ohio State University, College of Education, School of teaching and learning.	Pretest equivalency not established/ documented; No adequate control group

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Recovery	Rodgers, E., Gomez-Bellenge, F., Wang, C., Schulz, M. (April 2005). *Predicting the literary achievement of struggling readers: Does intervening early make a difference?* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec.	Pretest scores > 1/2 SD apart
Reading Recovery	Schmitt, M. (2001). The development of children's strategic processing in Reading Recovery. *Reading Psychology, 22*, 129-151.	Pretest equivalency not established/ documented; Demographic differences > 1/2 SD apart
Reading Recovery	Schmitt, M. (2003). Metacognitive strategy knowledge: Comparison of former Reading Recovery children and their current classmates. *Literacy Teaching and Learning,* 7(1-2), 57-76.	No adequate control group; Pretest equivalency not established/ documented
Reading Recovery	Schmitt, M., & Gregory, A. (2001, June). The impact of early intervention: Where are the children now? Paper presented at the meeting of the National Reading Conference, San Antonio, TX.	No adequate control group
Reading Recovery	Schmitt, M.C., Gregory, A.E. (2005). The impact of early literacy intervention: Where are the children now? *Literacy Teaching and Learning 10*(1), 1-20.	No adequate control group
Reading Recovery	Schotanus, H. (1991). Reading Recovery pilot project (Laws 1989: 301): Report of results and effectiveness. Concord, NH: New Hampshire State Department of Education. (ERIC No. ED363859).	No adequate control group
Reading Recovery	Schotanus, H., Chase, C. Fontaine, A., & Tilton, S. (1993). Report of results and effectiveness, Reading Recovery Program, implementation year three. New Hampshire Department of Education. (ERIC No. ED405573).	Pretest equivalency not established/ documented
Reading Recovery	Shanahan, T., & Barr, R. (1995). Reading Recovery: An Independent Evaluation of the Effects of an Early Instructional Intervention for At-Risk Learners. *Reading Research Quarterly, 30*(4), 958–996.	Outside of age/grade parameters
Reading Recovery	Shoulders, M.D. (2004). *The long-term effectiveness of the Reading Recovery program.* Unpublished doctoral dissertation, Tennessee State University.	Pretest scores > 1/2 SD apart

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reading Recovery	Smith, N. (1994). Reading Recovery data and observations from one Illinois site (part one). *Illinois Reading Journal 22*(2), 7-27.	No adequate control group
Reading Recovery	Smith, P.E. (1994). Reading Recovery and children with English as a second language. *New Zealand Journal of Educational Studies, 29*(2), 141-155.	No control group; Pretest equivalency not established/ documented
Reading Recovery	Spector, J.E. & Moore, P. (2003). Does phonological processing distinguish between students who are more or less responsive to Reading Recovery? *Literacy Teaching and Learning (8)*2, 1-25.	Pretest equivalency not established/ documented; Inadequate outcome measure
Reading Recovery	Wang, Y., & Johnstone, W. (1997, March). Evaluation of Reading Recovery Program. Paper presented at the meeting of the American Educational Research Association, Chicago.	No adequate control group
Reading Recovery	Yukish, J., & Fraas, J. (1997). Success of old order Amish children in a strategy-oriented program for children at-risk of failure in reading. In S. Swartz & A. Klein (Eds.), Research in Reading Recovery (pp. 39-51). Portsmouth, NH: Heineman.	No control group
Reading Recovery	Zielinski, Linda Alice (1997) The long term effectiveness of Reading Recovery in a small, rural school district. Ed.D. dissertation, Saint Louis University, United States -- Missouri. Retrieved September 5, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 9822883).	No adequate control group; Pretest scores > 1/2 SD apart
Reading Recovery	Daniel, A.F. (2007). Reading Recovery: An evaluation of one school district's academic outcomes. M.S. dissertation, University of Arkansas, United States -- Arkansas. Retrieved August 17, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 1442357).	Outside of age/grade parameters
Reading Recovery	Grehan, A., Ross, S., Harrison, L. & Nunnery, J. (2006). Evaluation of Reading Recovery In the Little Rock School District Technical Report. Center for Research in Educational Policy, the University of Memphis.	No control group
Reading Recovery	Bursiek, M.A. (1993). *Literacy interventions for low-achieving first graders.* Unpublished doctoral dissertation, Univeristy of Colorado and Boulder.	No adequate control group

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

reading recovery	Potter, W. (2007). An analysis of the achievement gap of discontinued reading recovery students: A longitudinal study of reading recovery students. Unpublished doctoral dissertation, The University of Nebraska - Lincoln, Nebraska.	No control group
Reading Recovery, DI (Horizons Fast Track)	Kahl, K.M. (2005). *Comparing outcomes of two early reading interventions: Reading Recovery and direct instruction.* Unpublished doctoral dissertation, Widener University.	Insufficient sample size
reading street	Wilkerson, S.B., Shannon, L.C., & Herman, T.L. (2006). *An efficacy study on Scott Foresman's Reading Street Program: Year one report.* Magnolia Consulting.	Outside of age/grade parameters; Pretest equivalency not established/ documented
reading street	Wilkerson, S.B., Shannon, L.C., & Herman, T.L. (2007). *An efficacy study on Scott Foresman's Reading Street Program: Year two report.* Magnolia Consulting.	insufficient information
Reading Success	Benson, K.A., Marchand-Martella, N.E., Martella, R.C., & Kolts, R.L. (2007). Assessing the Effects of the Reading Success Level B Program with Fifth-grade Students at a Title I Elementary School. Journal of Direct Instruction, Vol. 7, No. 1 – Winter 2007, pg 29-44	No control group
Reading to Read	Prestridge, C.C. (1996). Reading to Read and curriculum-based passages: Effects on student performance. Unpublished doctoral dissertation, University of Southern Mississippi.	Outside of age/grade parameters; Inadequate outcome measure
Reading Together	Jennings, C. (2004) *The Reading Together(TM) cross-age tutoring program and its effects on the English language proficiency and reading achievement of English language learners.* Unpublished doctoral dissertation, University of North Texas.	No control group; Outside of age/grade parameters
Reciprocal teaching	Diehl, Holly L. (2005) The effects of the reciprocal teaching framework on strategy acquisition of fourth-grade struggling readers. Ed.D. dissertation, West Virginia University	Insufficient sample size
Reciprocal Teaching	Johnson-Glenberg, Mina C., "Training Reading Comprehension in Adequate Decoders/Poor Comprehenders: Verbal Versus Visual Strategies," Journal of Educational Psychology, Vol. 92, No. 4, 2000, p.772-782.	Pretest scores > 1/2 SD apart; No adequate control group; Insufficient sample size

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Reciprocal Teaching	Lysynchuk, L. M., Pressley, M., & Vye, N. J. (1990). Reciprocal teaching improves standardized reading-comprehension performance in poor comprehenders. *The Elementary School Journal, 90,* 469–484.	Outside of age/grade parameters; Insufficient sample size
Reciprocal Teaching	Lysynchuk, L., Pressley, M., & Vye, N. J. (1989, March). *Reciprocal Instruction Improves Standardized Reading Comprehension Performance in Poor Grade-School Comprehenders.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.	Insufficient sample size
Repeated reading vs. continuous reading	O'Connor, R., White, A., & Swanson, H. (2007, September). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. Exceptional Children, 74(1), 31-46.	No adequate control group; Insufficient sample size
Repeated readings	Homan, S., Klesius, J., & Hite, C. (1993). Effects of repeated readings and nonrepetitive strategies on students' fluency and comprehension. *Journal of Educational Research, 87,* 94–99.	Insufficient sample size
Repeated Readings	Steventon, C., & Frederick, L. (2003). The effects of repeated readings on student performance in the corrective reading program. Journal of Direct Instruction, 3, 17-27.	Insufficient sample size
Restatement training	Jenkins, J., Heliotis, J., Stein, M., & Haynes, M. (1987). Improving reading comprehension by using paragraph restatements. Exceptional Children, 54, 54-59.	Duration < 12 weeks; Inadequate outcome measure
Rigby	Wilkerson, S.B. (August, 2004). A study of the effectiveness of Harcourt Achieve's Rigby Literacy Program: Final evaluation report. McRel.	Insufficient sample size
Road to the Code	Loughlin, J.E. (2003). *Early indentification and intervemtion with Kindergartners at risk for reading failure: A district-wide prevention program using a multiple gating approach.* Unpublished doctoral dissertation, University of Massachusetts - Amherst.	Inadequate outcome measure
RR vs Eagle Reading	Angeletti, S. (2000). Eagle Reading: A Comparison of a Small Group Reading Intervention with Reading Recovery, a One-on-one Tutorial. Unpublished doctoral dissertation, University of Georgia.	Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

SAIL	Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology, 88*(1), 18-37.	No control group; Pretest equivalency not established/ documented
Self questioning and story mapping	Taylor, L.K., Alber, S.R., & Walder, D.W. (2002). The comparative effects of a modified self-questioning strategy and story mapping on the reading comprehension of elementary school students with learning disabilities. *Journal of Behavioural Education, 11,* 69-87.	No control group; Pretest equivalency not established/ documented; Insufficient sample size
semantic relations (SR), or morphological awareness (MA)	Filippini, Alexis Louise (2007) Effects of a vocabulary-added instructional intervention for at-risk English learners: Is efficient reading instruction more effective? Ph.D. dissertation, University of California, Santa Barbara, United States -- California. Retrieved January 12, 2009, from Dissertations & Theses: Full Text database. (Publication No. AAT 3274436).	Duration < 12 weeks; Insufficient sample size
SFA	Koh, M., & Robertson, J. (2003). School reform models and special education. Education and Urban Society, 35 (4), 421-442.	No control group
SFA	Smith-Davis, Stacey L. (2007) Does Success For All impact reading achievement of students with learning disabilities. Ed.D. dissertation, University of Central Florida, United States -- Florida. Retrieved January 15, 2008, from ProQuest Digital Dissertations database. (Publication No. AAT 3276390).	Pretest scores > 1/2 SD apart; Inadequate outcome measure
slowing speech rate of instructors in PA training	Segers, E. & Verhoeven, L. (2004). Computer-Supported Phonological Awareness Intervention for Kindergarten Children with Specific Language Impairment. *Language, Speech, and Hearing Services in Schools, 35*, 229-239.	No control group; Insufficient sample size
SOAR	Bigby, G. (2008). The effects of instructional support programs on student achievement in reading. Unpublished doctoral dissertation, University of South Carolina.	No control group
Spatial Skills Training	Crano, W. (1991). Facilitating Reading Comprehension Through Spatial Skills Training. Journal of Experimental Education, 59(2), 113-127.	Outside of age/grade parameters

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Spell Read P.A.T.	Rashotte, C., MacPhee, K., & Torgesen, J. (2001). The Effectiveness of a Group Reading Instruction Program with Poor Readers in Multiple Grades. *Learning Disability Quarterly, 24* (2), 119-134.	No control group; Insufficient sample size
SPIRE	SPIRE Effectiveness Report	No control group; Outside of age/grade parameters
Spire Phonics Program	O'Donnell, M. (2001). Do intensive phonics programs help struggling readers?. *The New England Reading Association Journal*, *37*(2), 4-10.	Pretest equivalency not established/ documented; Duration < 12 weeks
SRA Open Court Reading	Cothran, J. (2006). Efficacy of a selection for reading interventions for low socioeconomic African-American students by ability and grade levels K-3. Unpublished doctoral dissertation, Tennessee Technological University	Pretest equivalency not established/ documented
STeps into Reading (STIR) + fluency	Ebaugh, J.C. (2000). *The effects of fluency instruction on the literacy development of at-risk first graders.* Unpublished doctoral dissertation, Forgham University.	Insufficient sample size
story grammar strategy	Rooney, J. (1997). The effects of story grammar strategy training on the story comprehension, self-efficacy and attributions of learning disabled students. *Dissertation Abstracts International, 58*(50A), 1642.	No control group; Insufficient sample size
Story mapping	Gardhill, M., & Jitendra, A. (1999). Advanced story map instruction: Effects on the readin gcomprehension of students withlearning disabilities. Journal of Special Education, 33, 2-17, 28.	Outside of age/grade parameters; Insufficient sample size
strategy training	Johnson, L., Graham, S., & Harris, K.R. (1997). The effects of goal setting and self-instructions on learning a comprehension strategy: A study with students with learning disabilities. *Journal of Learning Disabilities, 30,* 80-91.	Inadequate outcome measure
strategy training	Tam, K. , Heward, W. , & Heng, M. (2006). A Reading Instruction Intervention Program for English-Language Learners Who Are Struggling Readers. The Journal of Special Education, 40(2), 79-93.	Insufficient sample size
strategy training	Wilkins, S. (2007). Teaching Expository Text Strategies to Improve Reading Comprehension in Low Readers. Dissertation Abstracts International Section A: Humanities and Social Sciences, 68(6), 2409.	No control group

Success for All	Atkinson, C. (1998). An analysis of the impact of "Success for All" on reading, attendance, and academic self-efficacy with at-risk elementary school students. Dissertation Abstracts International, 59 (10), 3699A. (UMI No. 9905180).	Pretest equivalency not established/ documented
Success in Reading and Writing vs HBM	Lindsey, M.M. (1998). A comprehensive evaluation of an integrated reading and language arts curriculum, with attentionto the experiences of low achieving children. Unpublished doctoral dissertation, University of Oregon.	Insufficient sample size
Summarization and Self-Monitoring	Jitendra, A., Hoppes, M., & Xin, Y. (2000). Enhancing main idea comprehension for students with learning problems: The role of a summarization strategy and self-monitoring instruction. Journal of Special Education, 34, 127-139.	Pretest equivalency not established/ documented
Summer Reading Academy (used Soar to Success)	Durand, B.A.C. (2002). *The effect of the Summer Reading Academy on the reading achievement of struggling third grade readers.* Unpublished doctoral dissertation, University of Houston.	No adequate control group
SWELL (Schoolwide Early Language and Literacy), Reading Recovery	Center, Y., Freeman, L., & Robertson, G. (2001). The relative effect of a code-oriented and a meaning-oriented early literacy program on regular and low Progress Australian students in year 1 classrooms which implement Reading Recovery. *International Journal of Disability, Development and Education, 48* (2), 207-232.	Pretest equivalency not established/ documented
synthetic phonics (orton-gillingham)	Foorman, B.R., Francis, D.J., Winikates, D., Mehta, P., Schatschneider, C., & Fletcher, J.M. (1997). Early interventions for children with reading disabilities. Scientific Studies of Reading, 1, 255-276	Pretest scores > 1/2 SD apart
TAILS - tutoring	Al Otaiba, S., Schatschneider, C., & Silverman, E. (2005). Tutor-Assisted Intensive Learning Strategies in Kindergarten: How Much Is Enough? Exceptionality, 13(4), 195-208.	Outside age/grade parameters
text comprehension strategy	Van Den Bos, K.P., Brand-Bruwel, S., & Aarnoutse, C.A.J. (1998). Text comprehension strategy instruction with poor readers. *Reading and Writing: An Interdisciplinary Journal, 10,* 471-498.	Inadequate outcome measure

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

The Good Readers Club	Friesen, J.B. (2006). *An evaluation of the Good Readers' Club: An early literacy intervention program.* Unpublished master's thesis, Lakehead University.	Insufficient sample size
Time for Reading (tutoring)	Elliott, J., Arthurs, J., & Williams, R. (2000). Volunteer Support in the Primary Classroom: The Long-Term Impact of One Initiative upon Children's Reading Performance. British Educational Research Journal, 26 (2), 227-244.	No control group; Pretest equivalency not established/ documented
Total class peer tutoring	Kourea, L., Cartledge, G., Musti-Rao, S. (2007). Improving the reading skills of urban elementary students through total class peer tutoring, Remedial and Special Education, 28(2), 95-107	No control group; Pretest equivalency not established/ documented; Insufficient sample size
Tugmate (cross-age tutoring)	Barbetta, P. M., & Miller, A. D. (1991, February). Tugmate: A cross-age tutoring program to teach sight vocabulary. *Education & Treatment of Children* *14*(1), 19-38.	No control group
tutoring	Clark, N.M. (2007). *Investigating the relationship of in-class tutoring using focused reading strategies and the reading comprehension of struggling readers*. Unpublished doctoral dissertation, Capella University.	No control group
tutoring	Guy, M.M. (2001) *Effects of cross-age peer tutoring on the acquisition of early literacy skills in children attending kindergarten.* Unpublished doctoral dissertation, University of Nebraska - Lincoln.	Insufficient sample size
tutoring	Hitchcock, C.H., Prater, M.A. & Dowrick, P.W. (2004). Reading comprehension and fluency: Examining the effects of tutoring and video self-modeling on first-grade students with reading difficulties. Learning Disability Quarterly, 27, 89-103.	Inadequate outcome measure
Tutoring	Linan-Thompson, S., Vaughn, S., Hickman-Davis, P., & Kouzekanani, K. (2003). Effectiveness of supplemental reading instruction for English language learners with reading difficulties. *Elementary School Journal, 103*(3), 221-238.	No control group
tutoring (literacy support program, title 1 program)	Spaulding, C. (2006). Early literacy intervention as an alternative approach to instruction. Unpublished doctoral dissertation, University of Massachusetts.	Pretest equivalency not established/ documented

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

tutoring (stepping stones to literacy)	Benner, G. J. (2003). *An investigation of the effects of an intensive early literacy support program on the phonological processing skills of kindergarten children at-risk of emotional and behavioral disorders*. Unpublished doctoral dissertation, University of Nebraska.	Insufficient sample size
tutoring with Read Well or Read Naturally	Denton, C. A. (2000). The efficacy of two English reading interventions in a bilingual education program. *Dissertation Abstracts International,* 61(11), 4325A. (UMI No. 9994233)	Insufficient sample size
tutoring/ mentoring initiatives	Edwards, H.M. (2000). *The effects of tutorial and mentoring initiatives employed by military/school partnerships on selected improvement variables for at-risk elementary student in Bexar County, Texas.* Unpublished doctoral dissertation, Texas A&M University.	Pretest equivalency not established/ documented
tutoring/early intervention program	Dickinson, R.E. (2005). An assessment of the Early Intervention Program for reading for third-grade students at risk of failing the Georgia Criterion Referenced Competency Test in Monroe, Georgia. Ph.D. dissertation, Capella University	No control group; Pretest scores > 1/2 SD apart
tweaked version of Reading Recovery	Swain, A.M. (1998). An evaluation of an instructional intervention program based on Clay's Reading Recovery Program for elementary school students. Unpublished doctoral dissertation, Memorial University of newfoundland.	No adequate control group
Use of self generated main idea questions	Lubliner, S. (2004, February). Help for struggling upper-grade elementary readers. Reading Teacher, 57(5), 430-438.	Insufficient sample size
Various CAI	Wolf, Ann W. (2006) Using technology with learning-disabled readers: A meta-analysis. Ph.D. dissertation, Nova Southeastern University, United States -- Florida. Retrieved October 4, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3213866).	No control group; Pretest equivalency not established/ documented
video technology	Xin, J.F. & Rieth, H. (2001). Video-assisted vocabulary instruction for elementary school students with learning disabilities. *Information Technology in Childhood Education Annual,* 87-103.	Pretest equivalency not established/ documented; Insufficient sample size

Volunteer vs teacher tutors	Bell, Kathryn M. (2001) Partnerships in literacy tutoring: Using community volunteers to provide one-to-one tutoring to struggling readers in third, fourth, and fifth grade. Ph.D. dissertation, University of Pittsburgh	Insufficient sample size
WAT vs PAT	O'Shaughnessy, T. E. & Swanson, H. L. (2000). A comparison of two reading intervention for children with reading disabilities. Journal of Learning Disabilities, 33, 257-277.	Duration < 12 weeks
Watching movies to accompany listening to stories	Sharp, D.L.M., Bransford, J.D., Goldman, S.R., Risko, V., Kinzer, C.K., & Vye, N.J. (1995). Dynamic visual support for story comprehension and mental model building by young, at-risk children. *Educational Technology Research and Development, 43,* 25-42.	Pretest equivalency not established/ documented
Waterford Early Reading Program	Cassady, J., & Smith, L. (2005). The impact of a structured integrated learning system on first grade students' reading gains. *Reading and Writing Quarterly, 21*(4), 361-376.	Insufficient sample size
Wilson Reading System	Dickson, S., & Bursuck, W. (1999). Implementing a Model for Preventing Reading Failure: A Report from the Field. *Learning Disabilities Research and Practice, 14*(4), 191-202.	No control group
Wilson Reading System	Wilson, B.A. & O'Connor, J.R. (1995). Effectiveness of the Wilson Reading System used in public school training. In McIntyre, C. and Pickering, J. (Eds.). *Clinical Studies of Multisensory Structured Language Education.* Salem, OR: International Multisensory Structured Language Education Council.	No control group
	Vadasy, P., & Sanders, E. (2008). Code-Oriented Instruction for Kindergarten Students at Risk for Reading Difficulties: A Replication and Comparison of Instructional Groupings. *Reading and Writing: An Interdisciplinary Journal, 21*(9), 929-963.	No adequate control group
	Duffy-Hester, Ann Marie (1999) The effects of a balanced, accelerated, and responsive literacy program on the reading growth of elementary school struggling readers, unpublished doctoral dissertation, University of Georgia.	Insufficient sample size
	Knapp, N. F., & Winsor, A. P. (1998). A reading apprenticeship for delayed primary readers. Reading Research and Instruction, 38, 13-29.	No control group; Insufficient sample size

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

	Marston, D., Deno, S. Kim, D., Diment, K. & Rogers, D. (1995). Comparison of Reading Intervention Approaches for Students with Mild Disabilities. *Exceptional Children, 62*(1), 20-37.	Insufficient sample size
	Slavin & Madden 1991	Pretest equivalency not established/documented

Appendix 2

Abbreviations

AA-African American

ADD-Auditory Discrimination in Depth

ANCOVA- Analysis of Covariance

BAS-British Ability Scale

CAI- Computer-Assisted Instruction

CAT- California Achievement Test

C-Control

CIRC- Cooperative Integrated Reading Composition

CTBS- Comprehensive Test of Basic Skills

CTOPP-Comprehensive Test of Phonological Processing

DIBELS- Dynamic Indicators of Basic Early Literacy Skills

DI-Direct Instruction

DISP- Direct Instruction Skills Plan

DORT-Durrell Oral Reading Test

DRP- Degrees of Reading Power

E-Experimental

EIR-Early Intervention in Reading

ELL- English Language Learner

ELS-Early Literacy Support

ERIC- Educational Resources Information Center

ES- Effect Size

FL-Free Lunch

GORT- Gray Oral Reading Test

GRADE- Group Reading Assessment and Diagnostic Examination

H-Hispanic

HLM-Hierarchical Linear Modeling

HOSTS-Help Our Students to Succeed

ITBS- Iowa Tests of Basic Skills

K-ABC-Kaufman Assessment Battery for Children

K-Kindergarten

L-Large

M- Matched

MANCOVA- Multivariate Analysis of Variance

MAT- Metropolitan Achievement Test

MC-Metacognitive

MEAP-Michigan Educational Assessment Program

MPH- Matched Post-Hoc

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

MRRT-Metropolitan Reading Readiness Test

n.s.-not-significant

NAEP- National Assessment of Educational Progress

NALT-Northwest Achievement Levels Test

N-Number

P-Phonetic

PALS-Peer-Assisted Learning Strategies

PIAT-Peabody Individualized Achievement Test

PPVT-Peabody Picture Vocabulary Test

PTR-Programmed Tutorial Reading

R- Randomized Experiment

RAILS-Reading and Integrated Literacy Strategies

RI-Reading Intervention

RQE- Randomized Quasi-Experiment

RR-Reading Recovery

RTI-Response to Intervention

SAT- Stanford Achievement Test

SD- Standard Deviation

SDRT- Stanford Diagnostic Reading Test

SFA-Success for All

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

SG-Small Group

SHIP- Schools and Homes in Partnership

SMART-Start Making a Reader Today

S-Small

TDI-Teacher-Directed Instruction

TERA-Test of Early Reading Ability

TOWRE- Test of Word Reading Efficiency

TPRI-Texas Primary Reading Inventory

TPRI-Texas Primary Reading Inventory

WISC- Wechsler Intelligence Scale for Children

WJ- Woodcock-Johnson

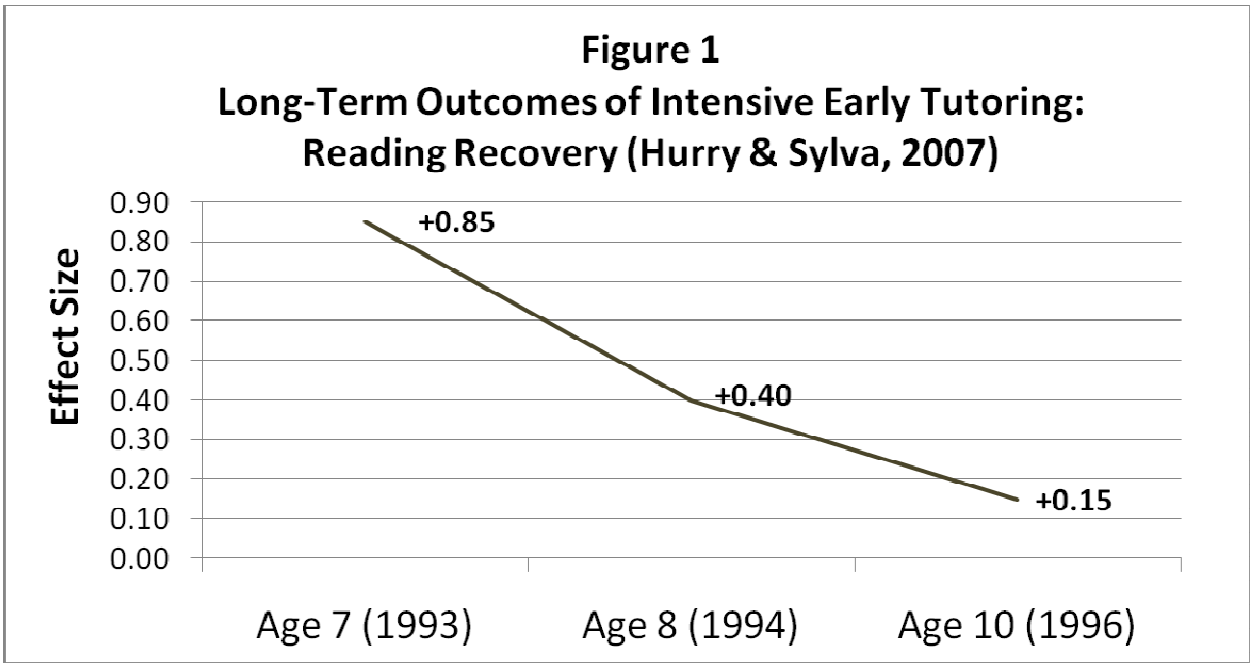
WRAPS-Word Reading and Phonics Skills

WRAT-Wide Range Achievement Test

WWC-What Works Clearinghouse

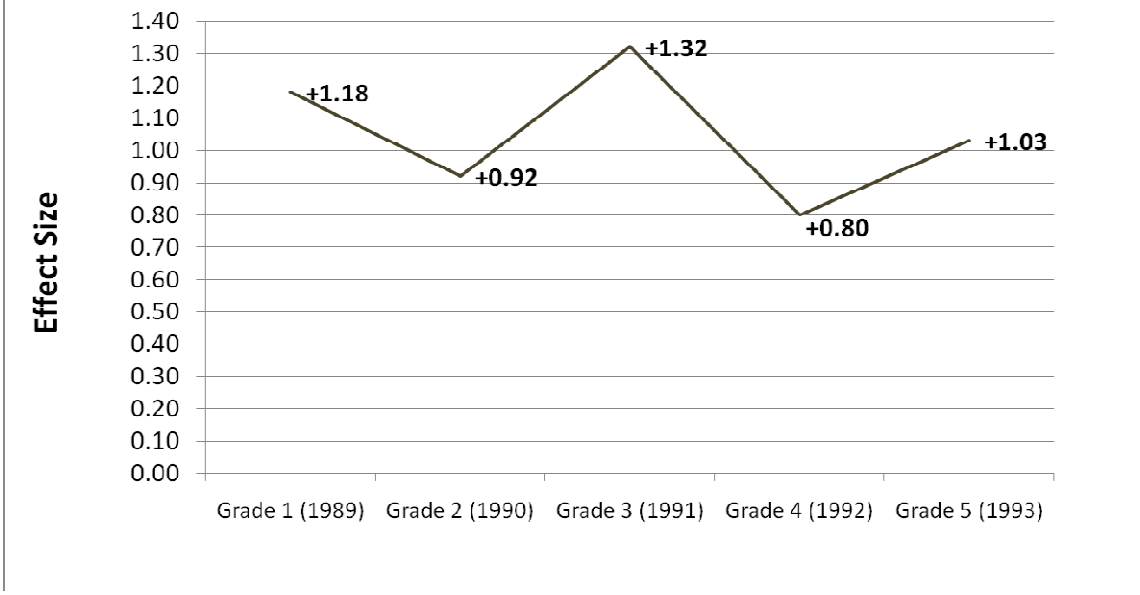
W-White

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.



The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Figure 2
Long-Term Outcomes of Early and Ongoing Intervention: Success for all (Slavin et al., 1993)



Note: Effect sizes are from Woodcock and Durrell scales in grades 1-3, Woodcock and Gray in grades 4-5.

Table 1
One-to-One Tutoring by Teachers

Study	Design Large/Small	Duration	N	Grade	Sample Characteristics	Posttest	Effect Sizes by Subgroup/ Measure	Overall Effect Size
Reading Recovery								
Pinnell et al. (1994)	Randomized (L)	1 year	33 schools 193 students (31E, 162C)	1	Urban, suburban, and rural schools in Ohio; 74%W, 26% AA, 65% FL	Gates MacGinitie		+0.19
Pinnell, DeFord, & Lyons (1988)	Randomized (S)	1 year	12 schools 187 students (126E, 26C)	1	Low-achieving students in Columbus, Ohio	CTBS		+0.52
						Comprehension	+0.55	
						Vocabulary	+0.48	
Curry, Griffith, & Williams (1995)	Matched (L)	1 year	553 students (268E, 285C)	1	High-poverty students in Austin, TX. 47%AA, 47%H, 6%W	ITBS		-0.16
Burroughs-Lange (2008)	Matched (S)	1 year	42 schools (21E, 21C) 234 students (87E,147C)	Ages 5-7	Low-achieving students in London, England with high proportions of ELLs and children receiving FL	BAS Word Reading	+0.87	+0.76
						WRAPS	+0.65	
Hurry & Sylva (2007)	Matched (S)	1 year	198 students (89E, 109C)	Ages 6-7	Primary schools in England 42% FL, 16% ELLs	BAS Word Reading	+0.84	+0.85
						Neale Prose Reading	+0.85	
Center, Wheldall, Freeman, Outhred, & McNaughton (1995)	Matched (S)	1 year	15 schools (10E, 5C) 56 students (22E, 34C)	Year 1	Urban schools in New South Wales	Neale Analysis of Reading Ability	+1.15	+0.86
						Passage Reading	+1.00	
						Cloze Test of Syntactic Awareness	+0.46	
						Word Attack Skills Test	+0.82	
Escamilla (1994)	Matched (S)	1 year	6 schools 46 students (23E, 23C)	1	Spanish-dominant students in Southern Arizona	Aprenda		+0.30
Huggins (1999)	Matched (S)	1 year	30 schools 122 students (70E, 52C)	1	High-poverty students in Detroit, MI	CAT		-0.09
						Comprehension	+0.03	
						Vocabulary	-0.15	
Other One-to-One Tutoring by Teachers								
Auditory Discrimination in Depth								
Torgesen, Wagner, & Rashotte (1997)	Randomized (S)	2 1/2 years	65 students (33E, 32C)	K-2	50%W, 49%AA	Woodcock Johnson		+0.90
						Word ID	+0.65	
						Word Attack	+1.02	
						Passage Comprehension	+0.39	
						Word Efficiency	+1.28	
						Non-Word Efficiency	+1.17	

Early Steps/ Howard Street Tutoring								
Morris, Tyner, & Perney (2000)	Matched (S)	1 year	11 schools (6E, 5C) 86 students (43E, 43C)	1	High-poverty AA schools in Tennessee	Woodcock Johnson adaptations		+0.86
						Word Attack	+0.92	
						Passage Comprehension	+0.80	
Brown, Morris, & Fields (2005) (Teacher tutors only)	Matched (S)	1 year	7 schools 59 students (17E, 42C)	2-6	Urban schools in the intermountain west 46% FL, 42% minority, 25% ELLs	Woodcock Johnson		+1.03
						Word Attack	+1.00	
						Passage Comprehension	+1.31	
						Word recognition	+0.81	
						Passage reading	+1.00	
Santa & Høien (1999)	Matched (S)	1 year	4 schools (2E, 2C) 49 students (23E, 26C)	1	Lower middle class schools in Montana	Woodcock Johnson		+1.04
						Word ID	+0.70	
						Word Attack	+1.28	
						Passage Comprehension	+1.14	
Reading Rescue								
Ehri et al. (2007) (Teacher tutors only)	Matched (S)	6 months	102 students (32E, 70C)	1	Spanish-dominant students in a large city 95%FL	Gates MacGinitie		+1.08
Reading with Phonology								
Hatcher, Hulme, & Ellis (1994)	Matched (S)	7 months	63 students: (32 E, 31C)	Ages 6-7	Schools in rural Northern England	BAS Word Reading	+0.40	+0.65
						Neale Analysis of Reading Ability		
						Accuracy	+0.62	
						Comprehension	+0.94	
Intensive Reading Remediation								
Blachman, Schatschneider, Fletcher, Francis, Clonan, Shaywitz, & Shaywitz (2004)	Randomized (S)	1 year	69 students (37E, 32C)	2, 3	School in 4 districts in upstate NY	Woodcock		+0.85
						Word ID	+0.99	
						Word Attack	+0.96	
						GORT Oral Reading Quotient	+0.75	
						Word Reading	+0.77	
						Word Reading Efficiency	+0.80	
Targeted Reading Intervention								
Vernon-Feagans, Amendum, Kainz, Ginsberg, & Bock (2009), Study 1	Randomized quasi-experimental (S)	1 year	6 schools 125 students (59E, 66C)	K, 1	Rural impoverished counties in the southeast; 80%FL, 49% AA, 33%W, 10% AI	Woodcock Johnson		+0.25
						Letter-Word ID	+0.24	
						Word Attack	+0.25	

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Vernon-Feagans, Amendum, Kainz, Ginsberg, & Bock (2009), Study 2	Randomized quasi-experimental (S)	1 year	4 schools 43 classrooms (26E, 17C) 151 students (97E, 54C)	K, 1	Rural Texas and New Mexico; 37%W, 26% AA, 35% Other	Woodcock Johnson		+0.34
						Word Attack	+0.27	
						Letter-Word ID	+0.24	
						Passage Comprehension	+0.50	
TEACH								
Mantzicopoulos et al. (1992)	Randomized (S)	2 years	116 students (59E, 57C)	1-2	Middle-class children in suburban San Francisco in lowest third of their class	SDRT-Comprehension	+0.10	+0.19
						SDRT-Phonetic Analysis	+0.09	
						Woodcock Word Attack	+0.30	
						K-ABC Reading/Decoding	+0.29	
Arnold et al. (1977)	Matched (S)	7 months	3 schools 63 students (23E, 40C)	1	2 inner city and 1 middle class school	WRAT-Reading		+0.34

Note: L=large study with at least 250 students; S=small study with less than 250 students; E=Experimental; C=Control; FL=Free/reduced-price lunch; W=White; AA=African American; H=Hispanic; AI=American Indian; ELL=English language learner; RR=Reading Recovery; CTBS=Comprehensive Test of Basic Skills; ITBS=Iowa Test of Basic Skills; CAT=California Achievement Test; ADD=Auditory Discrimination in Depth; BAS=British Ability Scales; WRAPS=Word Reading and Phonics Scale; GORT=Gray Oral Reading Test; SDRT=Stanford Diagnostic Reading Test; K-ABC=Kaufmann Assessment Battery for Children; WRAT=Wide Range Achievement Test.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Table 2
Tutoring by Paraprofessionals and Volunteers

Study	Design Large/Small	Duration	N	Grade	Sample Characteristics	Posttest	Effect Sizes by Subgroup/ Measure	Overall Effect Size
Tutoring by Paraprofessionals								
Sound Partners								
Jenkins, Peyton, Sanders, & Vadasy (2004)	Randomized (S)	1 year	11 schools 99 students (79E, 20 C)	1	Urban schools in the northwest	Woodcock Johnson		+0.69
						Word ID	+0.50	
						Word Attack	+0.77	
						Passage Comprehension	+0.81	
						TOWRE		
						Phonemic Decoding	+0.38	
						Sight Word Reading	+0.52	
						Bryant	+1.13	
						WRAT-Reading	+0.74	
Mooney (2003)	Randomized (S)	1 year	7 schools 47 students (28E, 19C)	1	Students at risk for emotional and behavioral disorders in midwestern city ; 55% FL, 68%W, 21%AA, 9%H	Woodcock Johnson		+0.34
						Basic Skills	+0.26	
						Comprehension	+0.22	
						DIBELS		
						Nonsense Word Fluency	+0.54	
						Oral Reading Fluency	+0.35	
Vadasy, Sanders, & Tudor (2007)	Randomized (S)	15 weeks	43 students (23E, 20C)	2-3	Urban schools in the northwest	Woodcock Word Attack and Word ID composite	+0.49	+0.52
						DIBELS-Fluency	+0.55	
Vadasy, Sanders, & Peyton (2005)	Matched (S)	1 year	57 students (38E, 19C)	1	Urban schools in the northwest	Woodcock Johnson		+0.71
						Word ID	+0.86	
						Word Attack	+1.06	
						Passage Comprehension	+0.66	
						TOWRE		
						Phonemic Decoding	+0.55	
						Sight Words	+0.56	
						Reading Rate	+0.33	
						Reading Accuracy	+0.66	
						WRAT	+1.02	
The Reading Connection								
Compton (1992)	Randomized (L)	1 semester	483 students (266E, 217C)	1	Kalamazoo, MI 53% minority	ITBS		+0.22
Start Making a Reader Today (SMART)								
Baker, Gersten, & Keating (2000)	Randomized (S)	2 years	6 schools 24 classrooms 84 students (43E, 41C)	1-2	Oregon	Woodcock Johnson		+0.50
						Word ID	+0.62	
						Passage Comprehension	+0.36	
						Word Comprehension	+0.46	
						Oral Reading Fluency	+0.54	

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Edmark Reading Program								
Mayfield (2000)	Randomized (S)	1 semester	3 schools 60 students (31E, 29C)	1	Low-achieving non-special-ed students in rural Louisiana	Woodcock Johnson		+0.23
						Passage Comprehension	+0.51	
						Letter-Word ID	+0.14	
						Word Attack	+0.03	
Wallach and Wallach								
Dorval, Wallach, & Wallach, (1978)	Matched (S)	1 year	58 students (20E, 38C)	1	Rural Roanoke Rapids, NC	CTBS Total Reading		+0.71
						Same-school controls	+0.66	
						Different-school controls	+0.77	
Programmed Tutorial Reading								
McGrady (1984)	Matched (S)	1 year	6 schools 69 students (35E, 34C)	4	Rural/suburban middle class schools in Indiana. Students below 37th percentile	ITBS Comprehension		+0.20
Reading Rescue								
Ehri et al. (2007) (Paraprofessional tutors)	Matched (S)	6 months	96 students (26E, 70C)	1	Spanish-dominant students in a large city; 95% FL	Gates MacGinitie		+0.89
Howard Street Tutoring								
Brown et al. (2005) (Paraprofessional tutors)	Matched (S)	1 year	63 students (21E, 42C)	1	Urban schools in the intermountain west. 46% FL; 42% minority, 25% ELLs	Woodcock Johnson		+0.55
						Word Attack	+0.07	
						Passage Comprehension	+0.95	
						Word Recognition	+0.60	
						Passage Reading	+0.58	
Tutoring by Volunteers								
Experience Corps								
Morrow-Howell, et al. (2009)	Randomized (L)	1 year	881 students (430E, 451C)	1-3	Schools in Boston, New York, Port Arthur, Texas 94%FL, 58%AA, 36%H, 24%ELL	Woodcock Johnson		+0.11
						Word Attack	+0.10	
						Passage Comprehension	+0.13	
Book Buddies								
Meier & Invernizzi (2001)	Randomized (S)	4 months	55 students (28E, 27C)	1	South Bronx, NY Students in lowest 25% of grade at end of K; 99% FL, 69%H, 30%AA	WRAT-Reading	+1.00	+0.89
						Words read correctly per minute	+0.78	
HOSTS								
Ramey (1991)	Matched (S)	1 year	238 students (18E, 220C)	2-5	Low-achieving students in Seattle	CAT-Reading		+0.05

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Other Volunteer Tutoring								
Ritter (2000)	Randomized (L)	1 year	11 schools 385 students (196E, 189C)	2-5	Philadelphia, PA. 87% FL, 96%AA	SAT-9		-0.10
Pullen et al. (2004)	Randomized (S)	3 months	10 schools 47 students (23E, 24C)	1	North-central Florida 53%W, 38%W, 58%FL	Woodcock Johnson		+0.52
						Letter ID	+0.23	
						Word Attack	+0.80	
Rimm-Kaufman, Kagan, & Byers (1999)	Randomized (S)	1 year	42 students (21E, 21C)	1	Cambridge, MA 29%AA, 26% Haitian Creole, 26% W, 60% FL	Clay Observational Survey		+0.27
						Reading Level	+0.35	
						Word Knowledge	+0.18	
Allor & McCathren (2004)	Cohort 1 Matched Cohort 2 Randomized	6 months	Cohort 1 8 schools 86 students (61E, 25C) Cohort 2 10 schools 157 students (76E, 81C)	1	Cohort 1: 100%FL, 100%AA Cohort 2: 94%FL, 96% AA	Cohort 1		+0.54
						Woodcock Johnson		
						Word ID	+0.59	
						Word Attack	+0.93	
						Passage Comprehension	+0.49	
						TOWRE Real Words	+0.41	
						TOWRE Non Words	+1.44	
						Cohort 2		
						Woodcock Johnson		
						Word ID	+0.11	
						Word Attack	+0.80	
						Passage Comprehension	-0.16	
						TOWRE Real Words	+0.14	
						TOWRE Non Words	+0.61	
						DIBELS Nonsense Word Fluency	+0.31	

Note: L=large study with at least 250 students; S=small study with less than 250 students; E=Experimental; C=Control; FL=Free/reduced-price lunch; W=White; AA=African American; H=Hispanic; ELL=English language learner; TOWRE=Test of Word Reading Efficiency; WRAT=Wide Range Achievement Test; DIBELS=Dynamic Indicators of Basic Early Literacy Skills; ITBS=Iowa Test of Basic Skills; CAT=California Achievement Test.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Table 3
Small Group Tutorials

Study	Design Large/Small	Duration	N	Grade	Sample Characteristics	Posttest	Effect Sizes by Subgroup/ Measure	Overall Effect Size
Corrective Reading: Decoding								
Torgesen et al. (2006, 2007)	Randomized (S)	1 year	16 schools 165 students (99E, 66C)	3 and 5	Schools around Pittsburgh 44% FL, 81% W, 19% AA	Average of Woodcock, TOWRE, AIMSweb, and GRADE		+0.16
						3rd grade	+0.22	
						5th grade	+0.09	
Hempenstall (2008)	Matched (S)	7 months	206 students (134E, 72C)	ages 8-11	Disadvantaged students in a Melbourne, Australia suburb	Woodcock Word Attack		+1.22
Spell Read								
Torgesen et al. (2006, 2007)	Randomized (S)	1 year	16 schools 196 students (115E, 81C)	3 and 5	Schools around Pittsburgh 44%FL, 69%W, 31%AA	Average of Woodcock, TOWRE, AIMSweb, and GRADE		+0.17
						3rd grade	+0.21	
						5th grade	+0.12	
Wilson Reading								
Torgesen et al. (2006, 2007)	Randomized (S)	1 year	16 schools 158 students (103E, 55C)	3 and 5	Schools around Pittsburgh, 48%FL, 56%W, 44%AA	Average of Woodcock, TOWRE, AIMSweb, and GRADE		+0.17
						3rd grade	+0.26	
						5th grade	+0.08	
Failure Free Reading								
Torgesen et al. (2006, 2007)	Randomized (S)	1 year	16 schools 219 students (113E, 104C)	3 and 5	Schools around Pittsburgh; 44%FL, 80%W, 20%AA	Average of Woodcock, TOWRE, AIMSweb, and GRADE		+0.05
						3rd grade	+0.10	
						5th grade	0.00	
Quick Reads								
Vadasy & Sanders (2008)	Randomized (S)	15 weeks	162 students (82E, 80C)	2-3	Urban Washington State; 30%W, 28%AA, 23%H, 16% Asian	Woodcock Word ID	+0.27	+0.22
						TOWRE Sight Word	+0.12	
						GORT Comprehension	+0.16	
						GORT Fluency	+0.30	
						DIBELS Fluency	+0..27	
Vadasy & Sanders (2008)	Randomized (S)	20 weeks	12 schools 119 students (54E, 65C)	4-5	Urban northwest; 40%AA, 25%W, 12%H, 9% Asian	Woodcock Johnson		+0.20
						Word ID	+0.33	
						Word Comprehension	+0.09	
						Passage Comprehension	+0.23	
						TOWRE Sight Word	+0.21	
						DIBELS Fluency	+0.16	

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Targeted Intervention								
Wang & Algozzine (2008)	Randomized (S)	1 year	6 schools (4E, 2C) 139 students (101E, 38C)	1	Urban schools, 80%FL, 89%AA or H	Woodcock		+0.19
						Word ID	+0.32	
						Word Attack	+0.43	
						Passage Comprehension	+0.13	
						DIBELS		
						Nonsense Word	-0.12	
Proactive Reading								
Mathes et al. (2005)	Randomized (S)	1 year	6 schools 162 students (80E 82C)	1	Large urban district in Texas	Woodcock Johnson		+0.49
						Word Attack	+0.76	
						Word ID	+0.51	
						Passage Comprehension	+0.21	
Responsive Reading								
Mathes et al. (2005)	Randomized (S)	1 year	165 students (83E, 82C)	1	Large urban district in Texas	Woodcock Johnson		+0.31
						Word Attack	+0.28	
						Word ID	+0.36	
						Passage Comprehension	+0.30	
New Heights Reading Program								
Lesnick (2006)	Randomized (S)	18 weeks	9 schools 59 classes (30E, 29C) 233 students (118E, 115C)	3 and 5	Philadelphia and suburban PA	DIBELS	-0.01	+0.02
						TOWRE	-0.02	
						DRP	+0.09	
Read Naturally								
Heistad (2005)	Matched (S)	1 year	2 schools 102 students (51E, 51C)	3 and 5	Minneapolis, MN	NALT	+0.21	+0.27
						Minnesota Comprehensive Assessments	+0.34	
Voyager Passport								
Ehri et al. (2007)	Matched (S)	6 months	122 students (52E, 70C)	1	Spanish-dominant students in a large city (FL=95%)	Gates MacGinitie		+0.31
						Decoding	+0.26	
						Comprehension	+0.36	
Empower Reading								
Lovett et al. (2000)	Randomized (S)	14 weeks	37 students (15 E, 22C)	1-4	Students with reading disabilties in Toronto	Woodcock Johnson		+0.71
						Word Attack	+1.23	
						Word ID	+0.45	
						Passage Comprehension	+0.49	
						WRAT	+0.68	
Lovett et al. (2008)	Randomized (S)	1 year	166 students (122E, 44C)	2-8	Below-grade-level students in Toronto	Woodcock Johnson		+0.25
						Word ID	+0.22	
						Word Attack	+0.37	
						Passage Comprehension	+0.02	
						WRAT Reading	+0.40	

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Schools and Homes in Partnership (SHIP)								
Gunn et al. (2005)	Randomized (S)	2 years	211 students (105E, 106C)	K-3	Rural districts in central Oregon, 62%H, 38%W	Woodcock Johnson		+0.34
						Letter Word ID	+0.31	
						Word Attack	+0.66	
						Vocabulary	+0.20	
						Comprehension	+0.29	
						DIBELS Oral Reading Fluency	+0.24	
Gottshall Small Group Phonics								
Gottshall (2007)	Randomized (S)	15 weeks	64 students (35E, 29C)	1	Boys in rural Nacogdoches, TX, 42%AA, 34%H, 20%W	TPRI		-0.10
						Graphophonemic Knowledge	-0.05	
						Reading Accuracy	-0.01	
						Reading Comprehension	-0.24	
Early Intervention in Reading								
Taylor, Short, Frye, & Shearer (1992)	Matched (S)	1 year	60 students (30E, 30C)	1	Students in a suburban Midwestern district	Gates MacGinitie		+0.82
Read, Write, and Type-Small Group								
Torgesen et al. (2009)	Randomized (S)	1 year	73 students (34E, 39C)	1	Florida schools	Woodcock Johnson		+0.36
						Word ID	+0.41	
						Word Attack	+0.59	
						Passage Comprehension	+0.33	
						TOWRE		
						Non-word	+0.26	
						Word	+0.22	
Lindamood Phoneme Sequence Program-Small Group								
Torgesen et al., 2009	Randomized (S)	1 year	74 students (35E, 39C)	1	Florida schools	Woodcock Johnson		+0.66
						Word ID	+0.63	
						Word Attack	+0.93	
						Passage Comprehension	+0.46	
						TOWRE		
						Non-word	+0.79	
						Word	+0.50	

Note: L=large study with at least 250 students; S=small study with less than 250 students; E=Experimental; C=Control; FL=Free/reduced-price lunch; W=White; AA=African American; H=Hispanic; ELL=English language learner; TOWRE=Test of Word Reading Efficiency; GRADE=Group Reading Assessment and Diagnostic Examination; GORT=Gray Oral Reading Test; DIBELS=Dynamic Indicators of Basic Early Literacy Skills; DRP=Degrees of Reading Power; NALT=Northwest Achievement Levels Test; WRAT=Wide Range Achievement Test; TPRI=Texas Primary Reading Inventory; CTOPP=Comprehensive Test of Phonological Processing.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Table 4
Classroom Instructional Process Approaches

Study	Design Large/Small	Duration	N	Grade	Sample Characteristics	Posttest	Effect Sizes by Subgroup/ Measure	Overall Effect Size
Cooperative Integrated Reading and Competition								
Stevens & Slavin (1995a)	Matched (S)	2 years	137 students (72E, 65C)	2-6	Special Education students in working-class suburb of Baltimore 9%FL, 95%W	CAT		+0.35
						Vocabulary	+0.37	
						Comprehension	+0.32	
Stevens & Slavin (1995b)	Matched (S)	2 years	76 students (40E, 36C)	2-6	Suburban Maryland 10%FL, 92%W	CAT		+0.81
						Comprehension	+0.85	
						Vocabulary	+0.76	
Bramlett (1994)	Matched (S)	1 year	149 students (82E, 67C)	3	Rural southern Ohio	CAT		+0.38
						Comprehension	+0.33	
						Total Reading	+0.33	
						Word Analysis	+0.56	
						Vocabulary	+0.30	
PALS								
Mathes & Babyak (2001)	Randomized quasi-experimental (S)	14 weeks	20 classes 56 students (27E, 29C)	1	Medium-sized district in Florida 50%W, 48% AA	Woodcock Johnson		+0.59
						Word Identification	+0.51	
						Word Attack	+0.89	
						Basic Skills	+0.71	
						Passage Comprehension	+0.23	
Mathes, Torgesen, & Allor (2001)	Matched (S)	16 weeks	24 classes (12E, 12C) 75 students (42E, 33C)	1	Southeastern district 47%W, 51%AA	TOWRE		+0.48
						Nonword Efficiency	+0.48	
						Word Efficiency	+0.34	
						Woodcock		
						Word Identification	+0.43	
						Word Attack	+0.58	
						Basic Skills	+0.55	
						Passage Comprehension	+0.50	
Mathes, Howard, Allen, & Fuchs (1998)	Matched (S)	16 weeks	20 classes 56 students (28E, 28C)	1	Southeastern urban district 37%W, 63%AA	Woodcock		+0.46
						Word Identification	+0.51	
						Word Attack	+0.69	
						Passage Comprehension	+0.19	
Mathes et al. (2003)	Matched (S)	16 weeks	15 teachers (7E, 8C) 59 students (31E, 28C)	1	Low achievers in a southeastern district 29%FL, 49%W, 42%AA	TOWRE		+0.43
						Non Word	+0.48	
						Word Efficiency	+0.13	
						Woodcock		
						Word ID	+0.41	
						Word Attack	+0.98	
						Passage Comprehension	+0.13	

Same Age Tutoring

Eldredge & Quinn (1988)	Randomized quasi-experiment (S)	1 year	5 schools 64 students (32E, 32C)	2	Middle class schools in Provo, UT	Gates MacGinitie		+1.55
						Comprehension	+1.33	
						Vocabulary	+1.77	
Reading and Integrated Literacy Strategies (RAILS)								
Stevens et al. (2008)	Matched (S)	1 year	3 schools 87 students (51E, 36C)	K-5	Low-achieving schools in a small city in central PA. 71%FL, 94%W	MAT		+0.49
						Special Education Students K-2	+0.47	
						Special Education Students 3-5	+0.50	
Contextually-Based Vocabulary Instruction								
Nelson & Stage (2007)	Randomized quasi-experiment (S)	3 months	73 students (41E, 32C)	3, 5	Midwestern schools 70%W, 24% H	Gates MacGinitie		+0.41
						Comprehension	+0.60	
						Vocabulary	+0.23	
Reading Styles								
Lashell (1986)	Matched (S)	1 year	90 students (47E, 43C)	2-6	Students with learning disabilities in rural Snohomish County, WA	Gray Oral Reading Test		+0.79
Brooks (1991)	Matched (S)	1 semester	2 schools 42 students (22E, 20C)	2-6	Title I students in Northwest Ohio	Spadafore Diagnostic Reading Test		+0.36
						Oral Reading	+0.21	
						Silent Reading	+0.51	
Direct Instruction								
Bowers (1972)	Randomized (S)	1 year	8 classes (4E, 4C) 123 students (60E, 63C)	1	Urban schools in Oklahoma 100%W	Gates MacGinitie		+0.25
						Comprehension	+0.15	
						Vocabulary	+0.35	
Davis (1995)	Matched (S)	1 year	2 schools 111 students (59E, 52C)	2	Title I students in southern Mississippi	SAT		+0.49
Project Read								
Greene (1991)	Matched (S)	1 year	224 students (112E, 112C)	1-3	Students below 25th percentile in Louisiana	CAT		+0.59
Precision Teaching								
Haring & Krug (1975)	Matched (S)	1 year	4 schools 54 students (24E, 30C)	Ages 9-12	Disadvantaged students with mental retardation; 70% AA	WRAT		+1.18

Note: L=large study with at least 250 students; S=small study with less than 250 students; E=Experimental; C=Control; FL=Free/reduced-price lunch; W=White; AA=African American; H=Hispanic; ELL=English language learner; CAT=California Achievement Test; MAT=Metropolitan Achievement Test; TOWRE=Test of Word Reading Efficiency; SAT=Scholastic Achievement Test, WRAT=Wide Range Achievement Test.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Table 5
Classroom Instructional Process with Tutoring (Success for All)

Study	Design	Duration	N	Grade	Sample	Posttest	Effect	Overall
Borman et al. (2007)	Randomized (L)	3 years	35 schools 703 students (362 E, 341 C)	K-2	Title I schools throughout the U.S., 72% FL, 57% AA, 31% W, 10% H	Woodcock		+0.28
						Word Identification	+0.22	
						Word Attack	+0.40	
						Passage Comprehension	+0.22	
Madden et al. (1993); Slavin et al. (1993)	Matched (L)	6 years	330 students (165E, 165C)	1-5	African American students in high-poverty schools in Baltimore, MD scoring in the lowest 25%	Average of Woodcock and DORT/Gray		+1.05
						1st grade	+1.18	
						2nd grade	+0.92	
						3rd grade	+1.32	
						4th grade	+0.80	
						5th grade	+1.03	
Ross, Nunnery, & Smith (1996)	Matched (S)	1 year	4 schools (2 E, 2 C) 138 students (42E, 96C)	1	Lowest performers in mostly Hispanic schools in Amphitheater District near Tucson, AZ	Woodcock		+0.54
						Word Identification	+0.44	
						Word Attack	+1.07	
						Passage Comprehension	+0.30	
						Durrell Oral Reading	+0.37	
Ross & Casey (1998b)	Matched (S)	2 years	8 schools 92 students (36E, 56C)	K-1	High-poverty schools in Ft. Wayne, IN; 75% FL, 45% minority	Average of Woodcock and Durrell		+0.35
Ross, Smith, & Casey (1994)	Matched (S)	3 years	2 schools 93 students (58E, 38C)	1-3	Students in Caldwell, ID	Average of Woodcock and Durrell		-0.20
Casey, Smith, & Ross (1994)	Matched (S)	1 year	4 schools 65 students (49E, 16C)	1	Schools in Memphis, TN	Average of Woodcock and Durrell		+0.54
Ross, Smith, Bond, Casey & Johnson (1993)	Matched (S)	3 years	4 schools 47 students (23E, 24C)	1-3	AA schools in Montgomery, AL	Average of Woodcock and Durrell		+1.16
Ross, Smith, & Casey (1995)	Matched (S)	4 years	4 schools 41 students (24E, 17C)	2-4	Students in Ft. Wayne, IN	Average of Woodcock and Durrell		+0.45
Smith, Ross, & Casey (1994)	Matched (S)	4 years	2 schools 38 students (21E, 17C)	1-4	African American schools in Memphis, TN	Average of Woodcock and Durrell		+1.14

Note: L=large study with at least 250 students; S=small study with less than 250 students; E=Experimental; C=Control; FL=Free/reduced-price lunch; W=White; AA=African American; H=Hispanic; ELL=English language learner; DORT= Durrell Oral Reading Test

Table 6
Instructional Technology

Study	Design Large/Small	Duration	N	Grade	Sample Characteristics	Posttest	Effect Sizes by Subgroup/ Measure	Overall Effect Size
Jostens (early form of Compass Learning)								
Becker (1994)	Randomized (S)	1 year	1 school 56 students	2-5	Low performing students in inner city Baltimore	CAT		+0.41
Sinkis (1993)	Matched (S)	1 year	422 students (228E, 194C)	3, 5, 6	Urban northeast	MAT Comprehension		+0.12
						Grade 3	+0.14	
						Grade 5	+0.22	
						Grade 6	-0.01	
Standish (1995)	Matched (S)	1 year	43 students (22E, 21C)	2	Title I students in suburban Delaware	MAT6 Reading		+0.55
Fast ForWord								
Rouse & Krueger (2004)	Randomized (L)	1 year	4 schools 454 students (237E, 217C)	3-6	High-poverty northeastern city schools, 59% FL, 66% H, 27% AA, 61% ELL	Connecticut Mastery Test		+0.05
Marion (2004)	Matched (S)	1 year	63 students (34E, 29C)	5-6	Schools in Appalachian TN. 52% FL, 100% W	Terra Nova		+0.15
Lexia								
Macaruso et al, (2006)	Matched (S)	1 year	10 schools 167 students (83E, 84C)	1	Title I students in schools in Boston, MA	Gates MacGinitie		+0.67
Other Supplemental CAI								
Dynarski et al. (2007); Campuzano et al. (2009) - Destination Reading - Waterford - Headsprout - Plan Focus -Academy of Reading	Randomized (L)	1 year	Cohort 1: 872 students (505E, 367C) Cohort 2: 232 Students (130E, 102C)	1	National. 49% FL, 44%W, 31%AA, 22%H	SAT-9		-0.07
						Cohort 1	+0.02	
						Cohort 2	-0.39	
Dynarski et al. (2007); Campuzano et al. (2009) - LeapTrack - Academy of Reading -Read 180 -Knowledge Box (cohort 1)	Randomized (L)	1 year	Cohort 1: 755 students (410E, 345C) Cohort 2: 95 students (52E, 43C)	4	National. 64% FL, 17%W, 57%AA, 23%H	SAT-10		+0.04
						Cohort 1	-0.01	
						Cohort 2	+0.48	
Becker (1994)	Randomized (S)	1 year	60 students	2-5	Schools in Baltimore, MD; 50% FL	CAT-Reading		+0.10

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.

Ramey (1991)	Matched (L)	1 year	282 students (62E, 220C)	2-5	Urban Washington State	CAT-Reading		+0.22
Bass, Ries, & Sharpe, (1986)	Matched (S)	1 year	2 schools (1E, 1C) 145 students (73 E, 72 C)	5-6	High-poverty schools in rural VA	SRA	+0.22	+0.18
						Virginia Basic Learning Skills Test	+0.13	
Chaing et. al (1978)	Matched (S)	1 year	4 schools 137 students (65E, 72C)	7-12 year-olds	Cupertino, CA, a middle class suburb of San Francisco	PIAT		+0.22
						Reading Recognition	+0.18	
						Reading Comrpehension	+0.26	
Roth & Beck (1987)	Matched (S)	1 year	6 classes (3E, 3C) 37 students (20E, 17C)	4	Low SES urban school	CAT Vocabulary	+0.98	+0.44
						CAT Reading Comprehension	-0.10	
Coomes (1985)	Matched (S)	1 year	4 schools 36 students (18E, 18C)	4	Middle class schools in TX. 90% W.	CTBS		+0.30

Note: L=large study with at least 250 students; S=small study with less than 250 students; E=Experimental; C=Control; FL=Free/reduced-price lunch; W=White; AA=African American; H=Hispanic; ELL=English language learner; CAT=California Achievement Test; MAT=Metropolitan Achievement Test; SAT=Scholastic Achievement Test; PIAT=Peabody Individual Achievement Test; CTBS=Comprehensive Test of Basic Skills.

The Best Evidence Encyclopedia is a free web site created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education (CDDRE) under funding from the Institute of Education Sciences, U.S. Department of Education.