Effects of Educational Technology Applications on Reading Outcomes for Struggling Readers:

A Best Evidence Synthesis

Alan C. K. Cheung
Johns Hopkins University

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

July, 2012

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 review examines the effectiveness of educational technology applications in improving the reading achievement of struggling readers in elementary schools. The review applies consistent inclusion standards to focus on studies that met high methodological standards. A total of 20 studies based on about 7,000 students in grades K-6 were included in the final analysis. Findings indicate that educational technology applications produced a positive but modest effect on the reading skills of struggling readers (ES=+0.14) in comparison to "business as usual" methods. Among four types of educational technology applications, small-group integrated applications such as *Read, Write, and Type (RWT)* and *Lindamood Phoneme Sequence Program (LIPS)* produced the largest effect sizes (ES=+0.32). These are tutorial educational technology applications that use small-group interaction tightly integrated with reading curriculum. Supplementary models, such as *Jostens* and *Lexia*, had a larger number of studies (N=12) and a more modest effect size (ES=+0.18). Comprehensive models *READ 180* and *Read About* (ES=+0.04) as well as *Fast ForWord* (ES=+0.06), did not produce meaningful positive effect sizes. However, the results of these two categories of programs should be interpreted with extreme caution due to the small number of studies involved. More studies are required to validate the effectiveness of all technology applications. Policy implications are discussed.

Keywords: Educational technology, reading achievement, elementary schools, struggling readers, meta-analysis, research review.

Despite substantial investments in reading instruction over the past two decades, far too many American students remain poor readers, and this has profound implications for these children and for the nation. According to the most recent National Assessment of Educational Progress (NAEP, 2011), fewer than half of fourth-grade students (42%) scored at or above the proficient level in reading. The results were more troubling for minorities and English language learners (ELLs). While 55% of White children achieved at or above the proficient level on NAEP, only 19% of African Americans, 21% of Hispanics, and 3% of ELLs scored at this level. Similar patterns were found for eighth graders' NAEP scores. Children who are not able to read well in the early grades tend to be at higher risk of performing poorly in later grades and other subjects, having emotional and behavioral problems, and dropping out of school (Lesnick et al., 2010). Concerted efforts have been made over the past 20 years among practitioners, researchers, and policy makers to develop policy and identify effective interventions to help struggling readers succeed in reading. For example, approaches such as improved initial teaching of reading, one-to-one tutoring, small-group tutorials, comprehensive school reform, and technology applications have been used for struggling readers in many schools across the country. Among these approaches, educational technology applications have become one of the most popular. With more struggling readers being integrated into general classrooms and the increasingly prevalent use of educational technology in today's classrooms, it is important that teachers, schools, and districts understand the effectiveness of various types of educational technology applications that are available to them to help improve the reading skills of struggling readers. The purpose of this review is to examine the effects of alternative types of educational technology applications for struggling readers, focusing on high-quality, rigorous evaluations.

Previous Reviews on Educational Technology Applications for Struggling Readers

Although research reviews on general interventions for struggling readers have been abundant (Boardman et al., 2008; Edmonds et al., 2009; Gersten et al., 2009; L. A. Hall, 2004; T. E. Hall, Hughes, & Filbert, 2000; Jitendra, Edwards, Sacks, & Jacobson, 2004; MacArthur, Ferretti, Okolo, & Cavalier, 2001; Okolo & Bouck, 2010; Scammacca et al., 2007; Slavin, Lake, Davis, & Madden, 2011; Stetter & Hughes, 2010; Vaughn et al., 2008; Wanzek, Wexler, Vaughn, & Ciullo, 2010), none of these reviews focused exclusively on the use of educational technology applications to enhance reading achievement for struggling readers in the elementary grades. In addition, many of these reviews included studies with serious deficiencies such as a lack of a control group, brief duration, and use of measures that were closely aligned with content taught to experimental but not control treatments. For example, in their review, Scammacca et al (2007) examined effective interventions for adolescent struggling readers in grades 4-12. A total of 31 studies were included, and the overall effect size was +0.95. However, over 60% of the studies included researcher-developed measures that were closely aligned with the treatment. The effect size was significantly lower (ES=+0.46) when studies with these questionable measures were excluded. Jitendra et al. (2004) carried out a review on vocabulary instruction for students with learning disabilities. Overall, results from the six CAI studies were mixed, with an overall effect size of +0.16. Many studies in this review had very brief durations, a few weeks or less. A review carried out by Stetter et al. (2010) examined the impacts of computer-assisted instruction on reading comprehension for struggling readers. The review covered three main areas: computerized versus printed reading materials, computerized readers to compensate for reading difficulties, and research on a variety of tools. The findings indicated that "some interventions have had at least a somewhat positive effect on student comprehension, while other efforts have shown less positive effects with more limited teacher involvement. (p. 3)" Like the two previous reviews, many of the included studies, as acknowledged by the authors, had "a weak or absent comparison group, insufficient information about the sample and outcome measures, as well as small sample sizes that made it difficult to generalize the findings."

The review by Slavin and his colleagues (2011) was the only one that applied consistent inclusion criteria to focus on studies that met high methodological standards. In their review, they identified a total of 97 studies that compared various approaches to helping struggling readers, including one-to-one tutoring, small-group tutorials, classroom process approaches (such as cooperative learning), comprehensive school reform, and technology. Fourteen out of the 97 studies were evaluations of educational technology applications in reading for elementary and secondary students. Their conclusion was that educational technology had a minimal impact on the reading achievement of struggling readers, with an overall sample size-weighted mean effect size of +0.09 across all studies. *Lexia* and *Jostens* were the only two programs that had promising effects. Since the publication of their review, several additional studies meeting high methodological standards have become available.

The purpose of this review is to examine the research up to the present on using educational technology applications to help teach struggling readers in elementary schools. Only studies that met our strict inclusion criteria were included. In addition to the overall effects, we were interested in exploring the differential impacts of moderator variables such as type of interventions, grade level, program intensity, research design, and recency of educational technology applications. It is important to note that this review does not attempt to determine the unique contribution of technology itself but rather the effectiveness of programs that incorporate use of educational technology. Technological components are often confounded with curriculum contents, instructional strategies, and other elements (Clark, 1983; Clark, 1985a; 1985b), making it difficult or impossible to identify the unique contriubtions of the technology.

Working Definition of Educational Technology

It is important to define the term "educational technology," since it has been used broadly in the literature. In this meta-analysis, educational technology is defined as a variety of electronic tools and applications that help deliver learning content and support the learning process, in this case for elementary struggling readers. Examples include computer-assisted instruction (CAI), integrated learning systems (ILS), and the use of video or embedded multimedia as components of reading instruction.

In this review, we identified four major types of educational technology applications: Traditional supplemental computer-assisted instruction (CAI), comprehensive models, small-group integrated supplemental programs, and *Fast ForWord* (a distinct approach emphasizing teaching of auditory discriminations). Supplemental CAI programs, such as *Destination Reading*, *Plato Focus*, *Waterford*, and *WICAT*, provide additional instruction at students' assessed levels of need to supplement traditional classroom instruction. Comprehensive models, including *READ 180* and *Read About,* use computer-assisted instruction along with non-computer activities as students' core reading approach. Small-group integrated models, including *Failure Free Reading*, *Read, Write, and Type (RWT), and Lindamood Phoneme Sequence Program (LIPS)*, are tutorial educational technology applications that use small-group interaction tightly integrated with the reading curriculum. *Fast ForWord (FFW)* supplements traditional CAI with software designed to retrain the brain to process information more effectively through a set of computer games that slow and magnify the acoustic changes in normal speech (Macaruso & Hook, 2001).

Review Methods

The review methods used here are similar to those used by Slavin, Lake, Chambers, Cheung, & Davis (2009), 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. Comprehensive Meta-analysis Software Version 2 (Borenstein, Hedges, Higgins, & Rothstein, 2005) was used to calculate effect sizes and to carry out various meta-analytical tests, such as Q statistics and sensitivity analyses. Similar to other research reviews, this study followed five key steps: 1. locating all possible studies; 2. screening potential studies for inclusion using preset criteria; 3. coding all qualified studies based on their methodological and substantive features; 4. calculating effect sizes for all qualified studies for further combined analyses; and 5. carrying out comprehensive statistical analyses covering both average effect sizes and the relationships between effect sizes and study features.

Literature Search Procedures

In an attempt to locate every study that could possibly meet the inclusion criteria, a literature search of articles written between 1980 and 2012 was carried out. Electronic searches were made of educational databases (e.g., JSTOR, ERIC, EBSCO, Psych INFO, Dissertation Abstracts), web-based repositories (e.g., Google Scholar), and educational technology publishers' websites, using different combinations of key words (e.g. educational technology, instructional technology, computer-assisted instruction, interactive whiteboards, multimedia, reading interventions, etc). We also conducted searches by program name. We attempted to contact producers and developers of educational technology programs to check whether they knew of studies that we had missed. References from other reviews of educational technology programs were further investigated. We also conducted searches of recent tables of contents of key reading journals for the past five years (2007 to 2012): *Educational Technology and Society, Computers and Education, American Educational Research Journal, Reading Research Quarterly, Journal of Educational Research, Journal of Adolescent & Adult Literacy, Journal of*

Educational Psychology, and *Reading and Writing.* Citations in the articles from these and other current sources were located.

Criteria for Inclusion

In order to be included in this review, studies had to meet the following inclusion criteria (see Slavin, 2008, for rationales).

1. The studies evaluated applications incorporating any type of educational technology, including computers, multimedia, interactive whiteboards, and other technology.

2. The studies involved students who were having difficulties learning to read in the elementary grades. These are defined as children with reading disabilities, students in the lowest 33% (or lower) of their classes, or any student receiving tutoring, Title I, special education, or other intensive services to prevent or remediate serious reading problems. Students identified only as low in socioeconomic status or as limited English proficient were not included unless they were also low in reading performance.

3. The studies compared students taught in classes using a given technology-assisted reading program to those in control classes using an or standard methods. If a study compared a given treatment to an alternative innovative treatment (rather than to a standard treatment), the different outcomes are noted in the text, but not included in the tables, which focus only on comparisons of experimental and control groups.

4. Studies could have taken place in any country, but the report had to be available in English.

5. 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. Studies in which students selected themselves into treatments (e.g., chose to attend an after-school program) or were specially selected into treatments (e.g., special education programs) were excluded unless experimental and control groups were designated after selections were made.

6. Pretest data had to be provided, unless studies used random assignment of at least 30 units (individuals, classes, or schools) and there were no indications of initial inequality. Studies with pretest differences of more than 50% of a standard deviation were excluded because, even with analyses of covariance, large pretest differences cannot be adequately controlled for as underlying distributions may be fundamentally different (Shadish, Cook, & Campbell, 2002).

7. 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 program (but unlikely to be emphasized in control groups) were excluded. Measures of skills that do not require interpretation of print, such as phonemic awareness, oral vocabulary, spelling, or writing, were excluded.

8. A minimum study duration of 12 weeks was required. This requirement was intended to focus the review on practical programs intended for use for the whole year, rather than brief investigations. Brief studies may not allow programs to show their full effect. On the other hand, brief studies often advantage experimental groups that focus on a particular set of objectives during a limited time period while control groups spread that topic over a longer period. Studies with brief treatment durations that measured outcomes over periods of more than 12 weeks were included, however, on the basis that if a brief treatment has lasting effects, it should be of interest to educators.

9. Studies had to have at least two teachers in each treatment group to avoid compounding of treatment effects with teacher effect.

10. Studied programs had to be replicable in realistic school settings. Studies providing experimental classes with extraordinary amounts of assistance (e.g., additional staff in each classroom to ensure proper implementation) that could not be provided in ordinary applications were excluded.

Both the first and second author examined each potential study independently according to these criteria. When disagreements arose, both authors reexamined the studies in question together and came to a final agreement.

Study Coding

To examine the relationship between effects and studies' methodological and substantive features, studies were coded. Methodological features included research design, sample size, and year of publication. Substantive features included type of education technology application, grade level, and program intensity. The study features were categorized in the following way:

1. Types of publication: Published and unpublished
2. Decade of publication: 1980s, 1990s, 2000s, and 2010s
3. Research design: Randomized design or quasi-experiment
4. Sample size: small (N <250) and large (N≥250)
5. Grade level: Primary (Grade K-3), and upper elementary (Grade 4-6)
6. Program types: Comprehensive models, small-group integrated programs, *Fast ForWord*, and supplemental programs.
7. Program intensity: low (≤75 minutes per week) and high (>75 minutes per week). These times included both time students were working with technology and time they were doing other closely associated off-line activities in a comprehensive, core program.

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 Size Calculation and Statistical Analyses

In general, effect sizes were computed as the difference between experimental and control individual student posttests after adjustment for pretests and other covariates, divided by the unadjusted posttest pooled standard deviation. Procedures described by Lipsey & Wilson (2001) and Sedlmeier & Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available, as when the only standard deviation presented was already adjusted for covariates or when only gain score standard deviations were available. If pretest and posttest means and standard deviations were presented but adjusted means were not, effect sizes for pretests were subtracted from effect sizes for posttests. Studies often reported more than one outcome measure. Since these outcome measures were not independent, we produced an overall average effect size for each study. After calculating individual effect sizes for all 24 qualifying studies, Comprehensive Meta-Analysis software was used to carry out all statistical analyses, such as Q statistics and overall effect sizes. Mean effect sizes across studies were weighted by sample sizes using a random-effects procedure.

Findings

Study Characteristics

Twenty studies* based on a total of about 7,000 students in grades K-6 met the inclusion standards. The main features and findings of the qualifying studies are summarized in Table 1. Of these, 11 were published articles and 9 unpublished reports. Only two were published in the 1980s, 4 in the 1990s, 7 in 2000s, and 7 in the 2010s. Thirteen studies used an experimental design, whereas the other 7 were quasi-experiments. The program intensity varied from 25 minutes to 450 minutes per week, with a mean of 150 minutes and a standard deviation of 112.

Overall Effects

The overall findings, summarized in Table 2, suggest that educational technology applications produced a positive but modest effect size (ES=+0.14) in comparison to traditional methods. Note that if we had used a fixed-effects weighting model, which gives greater weight to large studies, the mean effect size would have been only +0.08. The large Q value (Q_B=38.13, df=19, p<0.006) suggests that there is substantial variation in this collective set of studies. Both substantive and methodological variables will be used to model some of these variations.

==============

Insert Tables 1 and 2 here

==============

* A "study" was defined as a unique comparison of experimental and control treatments. Two articles reported on more than one treatment-control comparison.

In the following section, we will first present the results of substantive and methodological features of the entire set of studies followed by a more detailed description of each program and study.

Substantive Features of the Studies

Three substantive features of the studies were used to model variations in outcomes: type of interventions, grade level, and program intensity.

Types of interventions. Outcomes varied substantially according to types of interventions. The four studies of small-group integrated applications, *Failure Free Reading (FFR)*, *Read, Write and Type* (*RWT*), and *Lindamood Phoneme Sequence Program* (*LIPS*) produced the largest effect sizes (ES=+0.32). The 12 studies of supplemental programs, such as *Jostens* and *Lexia,* generated an effect size of +0.18. The mean effect size from the two qualifying studies of comprehensive models, represented by *READ 180* and *Read About*, was +0.04, and for the *Fast ForWord* program, two qualifying studies had an average effect size of +0.06. A narrative description of each study is presented later in this article.

Grade levels. Studies were organized in two grade levels: primary grades (K-3) and upper elementary (4^{th}-6^{th}). Two of the studies examined outcomes across grades but did not provide disaggregated data. Our findings indicate that the mean effect size for primary grades (ES=+0.36) was much larger than that for upper elementary grades (ES=+0.07). The mean effect size for the two-mixed grades studies was +0.25. The between-group difference was marginally significant (Q_B=4.66, df=2, p<0.09).

Program intensity. Program intensity was grouped into two categories: low intensity (the use of technology interventions, including any associated off-line activies, less than 75 minutes a week) and high intensity (more than 75 minutes a week). The effect sizes for low and high-intensity program were +0.08 and +0.19, respectively, but the difference was not statistically significant due to low power (Q_B=1.20, p<0.27).

Methodological Features of the Studies

Sensitivity Analysis

To check whether there are any outliers in this collection of studies that might skew the overall findings, a sensitivity analysis was performed (Borenstein, Hedges, Higgins, & Rothstein, 2009). The analysis indicated that the removal of any one effect size does not substantially affect the overall effect size.

Publication Bias

To assess the possible impact of publication bias, two statistical analyses were performed: Classic fail-safe N and Orwin's fail-safe N. The classic fail-safe N test estimates that in order to nullify the effect, a total of 121 studies with null results would be needed. Similarly, the Orwin's test indicates that the number of missing null studies to bring the existing overall mean effect size to 0.01 was 157. Given the results of these tests, there is no reason to believe that publication bias could account for the positive effect size.

As an additional test of the possibility of publication bias, we used a mixed-effects model to examine whether there was a significant difference between published journal articles and unpublished sources such as technical reports and dissertations. The overall effect sizes for published articles and unpublished publications were +0.25 and +0.04, respectively. The Q-value (Q_B=6.47, df=1, p<0.01) indicates substantial publication bias in this collection of studies. In other words, the overall effect sizes from the published journal articles were significantly larger than those in unpublished publications, a difference that is very typical in meta-analysis (Lipsey & Wilson, 2001).

Decade of Publication

The results indicated no trend toward more positive results in recent decades. The effect sizes for studies in the 80s, 90s, 00s, and 10s were +0.20, +0.18, +0.08, and +0.22, respectively. No statistically significant differences (Q_B=1.72, p<0.63) were found among different years of publications.

Methodological Features

In order to understand possible reasons for variations among these studies, we examined methodological features of the studies such as research design and sample size to see how they affect reading outcomes.

Research design. One potential source of variation is the presence of different research designs (e.g. Abrami & Bernard, 2006). In this collection of studies, we identified two main categories of research designs: randomized (N=13) and matched control studies (N=7). Randomized experiments were those in which students, classes, or schools were randomly assigned to conditions and the unit of analysis was at the level of the random assignment. Matched control studies were ones in which experimental and control groups were matched on key variables at pretest, before posttests were known. The average effect size for randomized experimental studies and matched control studies were +0.08 and +0.28, respectively. The mean effect size for quasi-experimental studies was about three times the size of that for randomized studies (p<0.06).

Sample size. Studies with small sample sizes typically produce much larger effect sizes than do large studies (Borenstein et al., 2009; Liao, 1999; Slavin & Smith, 2008). In this collection of studies, there were a total of 8 large studies with sample sizes of more than 250 students and 12 small studies with fewer than 250 students. A statistically significant difference was found between large studies and small studies (Q_B =11.84, df=1, p<0.00). The mean effect size for the 12 small studies (ES=+0.32) was much larger than that of large studies (ES=+0.04). This suggests the possibility that the small studies create unrealistic conditions of implementation, or possibly that small studies with null results are less likely to be reported.

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.

Design/size. After examining the effect of research design and sample sizes separately, we looked at the combined effect of these two moderator variables. The difference among the four groups was significant (Q_B =11.46 and p<0.00). Small matched control studies produced the largest effect size (ES=+0.34), followed by small randomized studies (ES=+0.28), large matched control studies (ES=+0.12), and large randomized studies (ES=+0.03). Within each research design, the effect sizes of small studies were much larger than those of large studies.

Narrative Descriptions of Qualifying Studies

To help the reader understand more about various types of interventions included in this collection of studies, we present a narrative description of these applications, and the context, design and findings of each study in the following section.

Supplemental CAI

Supplemental applications of computer-assisted instruction (CAI) are by far the most common applications of technology in reading. CAI usually consists of drill, practice, and self-tutorial materials with regular assessments and assignment of students to appropriate materials based on their unique performance levels. Students typically work on CAI in a lab or at the back of the class, in perhaps 2-3 half-hour sessions each week.

Jostens **(earlier version of** ***Compass Learning*****).** *Jostens* is an earlier form of an integrated learning system now called *Compass Learning*. The system is designed to provide an extensive set of assessments, which place students in an individualized instructional sequence. Students work individually on exercises designed to fill in gaps in their skills. *Jostens/Compass Learning* ILS programs are typically used 15-30 minutes per day, 2-5 days per week. Three qualifying studies examined the effectiveness of Jostens in the 1990s.

The first qualifying study was carried out by Sinkis (1993) to evaluate *Jostens* with Title I students in a pullout program in eight 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 (71E, 63C), MAT Reading Comprehension posttests adjusted for pretests had an effect size of +0.14 (n.s.). Corresponding effect sizes for fifth graders (83E, 61C) were +0.22 (n.s.), and for sixth graders (74E, 70C) the effect size was -0.01 (n.s.), for a mean across grades of ES=+0.12.

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 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).

Another small *Jostens* study was conducted by Standish (1995) with second graders in two suburban Delaware schools. The *Jostens* school had 4 teachers and 56 students, while 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.

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 (22E, 21C) was +0.55.

Across the three studies of Jostens, the weighted mean effect size was +0.19.

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. One hundred and sixty students typically participate in 2 to 4 20-30-minute sessions each week. Macaruso, Hook, & McCabe (2006) evaluated the *Lexia* programs in a year-long study in 10 first-grade classes in 5 Boston schools. One class in each school was assigned to the experimental group and another to the control group (n=83E, 84C). Over 50% of all students were eligible for free or reduced-price lunch. After adjusting for initial pretest differences, the mean effect size for Title 1 students was +0.67 (p<0.02) on the Gates-MacGinitie Reading Test.

Captain's Log (Braintrain®) **and** ***Destination Reading.*** Rabiner et al. (2010) carried out a randomized trial to examine the effectiveness of two computer-based interventions for students with attention difficulties: *Captain's Log (Braintrain®)* and *Destination Reading*. *Captain's Log* is a commercially available product that provides structured opportunities for exercising attention. *Destination Reading* is a popular computer-assisted program that targets five key skills: phonemic awareness, phonics, fluency, vocabulary, and comprehension. Seventy-seven first graders from five low-SES public schools in the southeastern United States were randomly assigned to one of three conditions: *Captain's Log* (n=25), *Destination Reading* (n=27), and control (n=25). Participants were well matched on pretests and demographics. Students in the *Captain's Log* group scored higher than the controls on two reading outcomes measures: DIBELS fluency (ES=+0.69) and WJ-III reading (ES=+0.10), with a median effect size of +0.40. On the other hand, the *Destination Reading* group scored only slightly higher than the controls: DIBELS fluency (ES=+0.10) and WJ-III reading (ES=+0.13), with a median effect size of +0.12.

Thinking Reader. A randomized study was conducted by Drummond et al. (2011) to examine the effectiveness of *Thinking Reader*, a software program designed to help improve the reading vocabulary and comprehension of students in Grade 5-8 using a reciprocal teaching approach (Brown & Palincsar, 1989). *Thinking Reader* is intended to be integrated with classroom discussion and peer interaction. A total of 2,407 sixth grade students (1286E, 1121C) with low SES backgrounds from 16 school districts in Connecticut, Massachusetts, and Rhode Island participated in a year-long study. At the end of the study, treatment students in the lowest achieving group (n=425) scored non-significantly higher than their counterparts in the control group (n=383) on both GMRT-Vocabulary and GMRT-Comprehension with effect sizes of +0.14 and +0.13, respectively.

Other Supplemental CAI

The two largest randomized studies of supplemental CAI applications were carried out by Dynarski, Agodini, Heaviside, Carey, & Campuzano, (2007) and Compuzano, Dynarski,

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.

Agodini, & Rall (2007) and Dynarski et al. (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 treatment effects (ES=+0.02, n.s.).

The same study evaluated four CAI programs at the fourth grade level: *Leapfrog, READ 180, Academy of Reading,* and *KnowledgeBox*, 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 2,265 total students (1231E, 1034C). N's for the lowest 33% were 410E, 345C. On the SAT10, there were no differences in the proportions of students scoring below the 33rd percentile (ES= -0.01).

Campuzano et al. (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, whose outcomes were reported by Dynarski et al. (2007). 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 year (19.7 hours) as in the first (42.6 hours). Controlling for pretests, the effect size for the proportion of children scoring below the 33rd percentile was -0.39. A weighted mean effect size for first graders across the two cohorts was -0.07.

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 (of teachers) 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 fourth graders across the two cohorts was +0.04.

Multiple CAI Programs

Coomes (1985) evaluated the use of a variety of drill and practice software programs (e.g. *Fundamental Punctuation Practice, Micro-Read, Spelling Program, Word Attack Program*). Participants were 112 students from 16 fourth grade classrooms in four schools in Texas. 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.

software chosen for the study was evaluated to coordinate with the basic fourth grade curriculum guide and the Macmillan basic series on each reading level. Students in the two treatment schools used the software programs for 30 minutes per week whereas students in the other two control schools received 30 minutes per week on the computer using mathematics software during the mathematics instructional period. At the end of the fourth grade year the students in all four schools were administered the CTBS. The effect size for the 36 low achievers (n=18E, 18C) was non-significant but positive (+0.30, n.s.).

In a small study in two Virginia Title I schools, Bass, Ries, & Sharpe (1986) evaluated the use of a variety of software programs (e.g. *Alpine Skier, Tank Tactics, and Big Door Deal*) in grades 5-6. Both groups received regular classroom instruction in reading and mathematics. Students in the treatment school using CAI for 25 minutes weekly as part of their Title 1 instruction (n=73) were compared to those in a matched school (n=72) using conventional, supplementary Title 1 instruction. 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 median effect size of +0.18.

Computer Networking Specialists (CNS)

Becker (1994) reported a randomized evaluation of an ILS program called *Computer Networking Specialist (CNS)*. CNS incorporates a variety of drill and practice and tutorial software from over 10 different publishers into its own system for managing assessment and assignment of tasks. The software was based on repetitive practice of isolated skills. 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. Students in the treatment group received three 30-minute CAI sessions weekly. 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).

Across 12 studies of supplemental CAI, the weighted mean effect size was +0.18 (p<.02).

Comprehensive Models

READ 180. *READ 180* is one of the most widely used approaches for adolescent struggling readers. The model is intended to serve as a comprehensive literacy intervention, which combines computer and non-computer instruction in the classroom, with the support of extensive professional development for teachers. In a typical *READ 180* classroom, students are provided with a daily 90-minute reading lesson in a group of no more than 15 students. The lesson consists of 20 minutes of whole-class teaching followed by three 20-minute rotation activities in groups of 5, including computer-assisted instruction in reading, modeled or independent reading, and small-group instruction with the teacher. The class then ends with a whole-group wrap-up for 10 minutes. Teachers are given materials and professional development to support instruction in reading strategies, comprehension, word study, and vocabulary (Davidson & Miller, 2002). Numerous *READ 180* studies have been conducted in the past decade. However, the majority of them were at the secondary level. Slavin, Lake, & Groff,

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.

(2008) found positive effects for *READ 180* in middle schools (with a weighted mean effect size of +0.24 across eight studies). Two recent randomized studies with struggling readers at the elementary level were included in this review.

The first qualifying *READ 180* study at the elementary level was carried out by Kim Samson, Fitzgerald, & Hartry (2010). Approximately 300 fourth to sixth graders who scored below proficiency on the Massachusetts Comprehensive Assessment System (MCAS) were randomly assigned to either a modified version of *READ 180* or to the district's regular after-school program. To fit the after-school program schedule, the modified *READ 180* was shortened to 60 minutes and included only three key components: individualized computer-assisted reading instruction, independent and modeled reading practice with leveled text, and teacher-directed reading lessons tailored to the reading level of small groups but without a teacher-led vocabulary session. Effects were only found on TOWRE reading fluency with fourth graders. No other significant effects were found on other measures or grades, with an overall effect size of +0.03.

The second qualifying study conducted by Kim and his colleagues (Kim, Capotosto, Hartry, & Fitzgerald, 2011) built on their previous work. Unlike the modified *READ 180* version used in the first study, the later study used a *READ 180 Enterprise* version that was designed to conclude with a teacher-directed whole-group wrap-up lesson to review key objectives. The control group was not given whole-group instruction, individualized computer-assisted reading instruction, or independent and modeled reading practice with leveled text, so it is important to note that the control group was receiving much less time in reading. Participants were 312 fourth to sixth graders who scored below proficiency on the Massachusetts Comprehensive Assessment System (MCAS) from four elementary schools in a midsize urban district in southeastern Massachusetts. Students within each grade and school were randomly assigned to either *READ 180 Enterprise* or the district after-school program. The treatment students outperformed the controls on SAT-10 vocabulary and reading comprehension with effect sizes of +0.23 and +0.31, respectively. No significant differences were found on DORF oral reading fluency (ES=+0.10). The overall effect size across three measures was +0.21.

Read About. A large-scale randomized study was conducted by James-Burdumy et al. (2009) to evaluate four reading comprehension interventions, including *Project CRISS, Read About, Read for Real*, and *Reading for Knowledge. Read About,* developed by Scholastic, was the only educational technology program in this study. Students in *Read About* are taught reading comprehension skills, vocabulary, and content knowledge through an adaptive computer program three times a week for 20 minutes. In addition, students use offline materials once per week for 20 minutes. Offline materials include whole-class or small-group lessons on comprehension skills, vocabulary strategies, text types, or writing skills. Students rotate among computer, teacher-led, and independent reading groups. Teacher materials include suggestions for English language learners and differentiated instruction. Over 2,600 fifth graders from low-SES schools participated in the study. The number of struggling readers (the bottom third of students) were 415 and 456 for the treatment and control groups, respectively. No significant differences were found between the treatment and the controls after the one-year study period. The overall median effect size across TOSCRF Composite Test Scores and GRADE scores was -

0.03. Similar results were found for the other three non-technology programs. As a group, the combined treatment group scored lower than the control group, with an effect size of -0.08 across the two measures.

Across three studies, the weighted effect size for comprehensive models was near zero (ES= +0.04, n.s.).

Small-Group Integrated Programs

Failure Free Reading. Torgesen et al (2007) carried out a large randomized study to examine the effectiveness of four widely used remedial reading instructional programs for struggling readers in 16 elementary schools: *Corrective Reading*, *Failure Free Reading*, *Spell READ P.A.T.*, and *Wilson Reading.* All four interventions delivered instruction to groups of three students pulled out of their regular classroom activities. *Failure Free Reading* was the only program that had a technology component. It combines computer-based lessons, workbook exercises, and teacher-led instruction to teach sight vocabulary, fluency, and comprehension. In addition, it was the only program that emphasized building students' vocabulary of sight words rather than phonemic decoding strategies. The participating schools were first randomly assigned to one of the four interventions. Within each school and grade, students were then randomly assigned to either the treatment or control condition. Within the *Failure Free Reading* schools, 51 third graders and 62 fifth graders were assigned to receive the treatment and 38 third graders and 66 fifth graders were in the control groups. Students were pre- and posttested on a battery, including the Woodcock, TOWRE, AIMSweb, and GRADE. The third graders in the treatment group outperformed the controls on four of the seven reading outcome measures, with a median effect size of +0.19. Among fifth graders, no significant differences were found between the treatment and control groups (ES=-0.05). The combined effect size for the two groups was +0.05. The other three non-technology programs produced very similar results. Combining all four treatment groups, the overall effect sizes for 3rd and 5th grades were +0.17 and +0.03, respectively.

Read, Write, and Type (RWT) and Lindamood Phoneme Sequence Program (LIPS). Torgesen, Wagner, Rashotte, Herron, & Lindamood (2010) examined the relative effectiveness of two computer-assisted instructional programs designed to improve reading skills for first grade students who were at risk for dyslexia: *Read, Write, and Type* (*RWT*) and *Lindamood Phoneme Sequence Program* (*LIPS*). Both programs were designed to provide explicit and systematic support for the development of phonemic awareness, phonemic decoding, and text reading accuracy. Unlike other supplemental CAI programs, these are wellintegrated with classroom instruction. Both treatment groups were taught in small groups in four 50-minute weekly sessions for one school year, October to May. The control group did not receive any small-group or computerized remedial instruction. While the focus of the *RWT* program was more on teaching students "the spellings of phonemes and using that knowledge to support spelling and writing activities," the *LIPS* program was designed to provide "powerful support for the development of oral motor awareness in support of early decoding (reading) and encoding (writing) activities." Participants were 112 first graders from two cohorts randomly assigned to one of the three conditions: *RWT* (n=34), *LIPS* (n=35), or control (n=39). Both the *RWT* and

LIPS groups performed better than the controls on all five Woodcock and TOWRE reading outcome measures with a combined median effect size of +0.36 and +0.66, respectively. There was no significant difference between *RWT* and *LIPS*, however.

Across three studies of small-group integrated programs, the overall effect size was +0.32 (p<.09). This was by far the largest mean effect size among the four categories of programs, but it is important to note that the larger positive effects were all from very small studies.

Studies of Variations in Small-Group Programs

We identified three additional studies of small-group integrated programs. All three examined the effectiveness of adding a technology component to an existing innovative program. Though they did not meet inclusion criteria for Table 1 because they lacked an untreated control group, their findings are important.

Adding embedded multimedia and computer-assisted tutoring to Success for All. Chambers et al. (2008) carried out a study that examined the effectiveness of a combination of two technology applications for teaching beginning reading: Embedded multimedia and computer-assisted tutoring. A total of 159 first graders (T=75, C=84) from two high-poverty *Success for All* schools were randomly assigned to technology or non-technology conditions in a year-long study. In the treatment condition, all students were instructed in reading using *Reading Reels*, a technology-enhanced version of *Success for All* with embedded multimedia, designed to enhance the effectiveness of the program by giving children compelling, memorable video demonstrations of letter sounds, sound-blending strategies, vocabulary, and comprehension strategies on interactive whiteboards. Students who qualified for tutoring also used *Alphie's Alley*, a computer-assisted tutoring program aligned with the SFA curriculum, with a human tutor for 20 minutes daily. The tutoring program had four key components: assessment, planning, computer activities, and just-in-time professional development. Students in the control group used the regular *Success for All* reading program, including 20-minute daily tutoring for low performing children, without the technology components. Approximately one third of the students (T=32, C=28) received tutoring in both conditions. At the end of the study, the tutored students in the treatment group significantly outperformed their counterparts in the control group on all five reading outcome measures: Letter Word ID (ES=+0.47), Word Attack (ES=+0.39), GORT Fluency (ES=+0.58), GORT Comprehension (ES=+1.02), and GORT Total (ES=+0.76). The overall effect size was +0.64.

Adding the Team Alphie small-group computer-assisted tutoring program to Success for All. Chambers, Slavin, Madden, Abrami, Logan, and Gifford (2011) compared the relative effects of a Tier II computer-assisted tutoring in small groups (*Team Alphie*) to one-to-one tutoring without technology. Over 300 first and second grade struggling readers were identified from 33 high-poverty Success for All schools in nine states to participate in the study. Schools were randomly assigned to implement either *Team Alphie* or regular (paper-and-pencil) SFA one-to-one tutoring for a year. *Team Alphie* was designed to create a small-group supplementary reading intervention for students and was closely linked to the core reading instruction of the SFA program. The key components of *Team Alphie* include cooperative learning, computer-assisted instruction, embedded multimedia, and tutoring. *Team Alphie* is similar to *Alphie's Alley* (Chambers et al, 2008) but is intended for use with groups of up to six children. Students in the control group received the non-technology tutoring long provided in Success for All schools in a one-to-one format. The idea was to see if computer-assisted tutoring to small groups could be as effective as one-to-one tutoring, thereby making tutoring more cost-effective. After adjusting for initial differences, first graders in the treatment group scored significantly higher than the controls on all three measures: Woodcock Letter Word Identification (ES=+0.17), Word Attack (ES=+0.21), and Passage Comprehension (ES=+0.15). However, no significant differences were found between the treatment and control group for second grade, with an overall mean effect size of +0.01. The findings of this study provided some evidence that a computer-assisted, small-group tutoring program may be at least as effective as one-to-one tutoring without technology.

Computer-Assisted Remedial Reading Intervention (CARRI) compared to non-technology tutoring. A 2-year randomized longitudinal study was conducted by Saine et al. (2010) in Finland to examine the effectiveness of three types of reading interventions: remedial reading intervention (RRI), computer-assisted remedial reading intervention (CARRI), and mainstream instruction. Participants were 166 seven-year-old Finnish children from a middle-class suburban area in the province of Western Finland. After initial screening assessments, only the lowest 33% of these children were offered remedial reading intervention. They were then randomly assigned to one of the two remedial interventions: RRI (N=25) and CARRI (N=25). The rest were assigned to mainstream instruction. Both RRI and CARRI groups received four weekly sessions of 45-minute remedial intervention over a period of 28 weeks in first grade. Students were divided into groups of five and each remedial intervention session had four segments: pre-reading activities, word segmentation, decoding and spelling, and vocabulary training. Students in the CARRI group used the same phonics-based remedial reading package as the RRI group with the exception that the treatment students received a 15-minute individual time with a computer-assisted application called GraphoGame in replacement of the non-technology pre-reading activities in the RRI group. The CAI application was designed specifically for children with learning disabilities and risk for dyslexia and provided structured drill and practice in pre-reading and reading skills. At the end of the first grade, the CARRI group outperformed the RRI Group on both Letter Knowledge (ES=+0.59) and Reading Fluency (ES=+0.51). In a follow-up test at the end of second grade, the treatment group continued to maintain the lead over RRI on reading fluency (ES=+0.67).

Fast *ForWord* (FFW). *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 children using 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.

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.

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

The one randomized study of *Fast ForWord* that met the 12-week duration criterion was an evaluation by Rouse & Krueger (2004), involving four schools in a Northeastern city. All schools were implementing *Success for All* (Slavin, Madden, Chambers, & Haxby, 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 was given just before and just after treatment, and thus did not meet the duration requirement of 12 weeks (They did not show any significant differences on reading outcomes in any case). 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.

The second qualifying study was conducted by Marion (2004) with a group of 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 (34E, 29C) scored non-significantly higher (ES=+0.15, n.s.).

The two studies of *Fast ForWord* had a weighted mean effect size near zero (ES=+0.06, n.s.).

Discussion

The purpose of this review was to examine the overall effectiveness of educational technology applications on reading outcomes for struggling readers. We identified a total of 20 high quality studies that met our inclusion criteria. Our findings indicate that educational technology applications had a modest impact on reading achievement of struggling readers, with an overall weighted mean effect size of +0.14. The effect size is similar to the effect size of +0.16 reported in a recent review carried out by Cheung & Slavin (in press) for all students in K-12 classrooms. Among the four types of educational technology applications, small-group integrated applications such as *Read, Write, and Type (RWT) and Lindamood Phoneme Sequence Program (LIPS)* produced the largest effect sizes (ES=+0.32), but these were mostly small studies, which tend to overstate program impacts. Supplementary models, such as Jostens, had a larger number of studies and a more modest effect size (ES=+0.16). Comprehensive models (ES=+0.04) and *Fast ForWord* (ES=+0.06) did not produce meaningful positive effect sizes. However, the results of these two categories of programs should be interpreted with extreme caution due to the small number of studies involved.

It is not surprising that the largest effects were found in the small-group integrated supplemental programs. Previous studies have found that small-group tutorials were effective for struggling readers (Slavin et al. 2011). Unlike traditional supplemental programs, these small-group integrated programs were tightly integrated with existing curriculum in small-group settings. For example, *RWT* and *LIPS* were designed to provide explicit and systematic support of the development of phonemic awareness, phonetic decoding, and text reading accuracy in small-group settings. The findings from these experimental studies provide additional evidence that small-group integrated supplemental programs have a greater impact on reading outcomes for struggling readers than traditional methods. Further, studies by Chambers et al. (2008), Chambers et al. (2011), and Saines et al. (2010) found that adding well-integrated technology to small-group and one-to-one tutoring models enhanced their effectiveness.

In contrast to reviews of applications in secondary schools, comprehensive models such as *READ 180* and *Read About* did not generate meaningful effects on reading outcomes for struggling readers in elementary schools. For example, in their review, Slavin et al. (2008) identified a total of 8 qualifying *READ 180* studies in middle schools and concluded that there was moderate evidence of a positive impact on reading comprehension (ES=+0.24). However, the results of the two qualifying *READ 180* studies in elementary schools were mixed. The first did not find any statistically significant differences between the treatment and control group in reading outcomes (ES=+0.03), whereas the second found more positive results (ES=+0.21). It is important to point out that unlike the full 90-minute version of *READ 180*, these two studies used a modified 60-minute version that was used in an after-school setting. As mentioned earlier, a typical 90-minute *READ 180* class includes 20 minutes of whole-class teacher-directed instruction of high-utility words that appear frequently across content areas and three 20-minute instructional activities designed to improve reading efficiency, reading comprehension and vocabulary, and oral reading fluency. In contrast, the modified 60-minute version included individualized computer-assisted reading activities, independent and modeled reading of leveled books, and teacher-directed lessons for small groups of children. The researchers suspected that "the absence of 30-minute whole group instruction may have limited vocabulary gains and the overall efficacy of the *READ 180* intervention." Though the second study also used a similar modified 60-minute version of *READ 180*, it included regular teacher-directed whole-group instruction and whole-group wrap-up. The results were more in line with those of previous studies. The findings of these two studies provide some suggestive evidence that teacher-directed whole-group activities that provide students with systematic and explicit instruction in vocabulary may be beneficial in improving reading comprehension. However, given the small number of studies involved, there is a clear need for more studies evaluating *READ 180* in elementary schools.

There is some evidence that technology applications for struggling readers may be more effective with younger children than older children (Kim et al., 2010; Torgesen, 2007). The 8 qualifying studies that took place in the primary grades had an overall effect size of +0.36 whereas the 10 studies carried out in the upper-elementary grades produced an effect size near zero (ES=+0.07). This finding provides some evidence that early intervention is essential for struggling readers.

It appears that high intensity programs had a bigger impact on struggling readers than low intensity programs. The effect sizes for low-intensity and high-intensity programs were +0.08 and +0.19, respectively. It is important to note that the majority of these high-intensity programs combined technology and non-technology components in their reading interventions. For example, *Read, Write, and Type (RWT) and Lindamood Phoneme Sequence Program (LIPS)* were designed to provide explicit and systematic support for reading development in the small-group setting. Unlike traditional CAI approaches, these programs were well integrated with classroom instruction. Computer and integrated non-computer activities were taught about 200 minutes per week or 50 minutes daily. These programs become core, daily activities for students, not supplements, and this may account for their apparent effectiveness.

In addition to these overall findings, several interesting findings also emerged from this review and warrant a brief mention. First, 13 out of the 20 qualifying studies (65%) used randomized experiments to evaluate program effectiveness. Compared to previous reviews, this percentage of randomized studies is surprisingly high. Importantly, we found a significant difference between experimental and quasi-experimental designs. Effect sizes were generally three times larger in quasi-experiments than in randomized experiments.

Consistent with previous findings (Cheung & Slavin, in press; Pearson, Ferdig, Blomeyer, & Moran, 2005; Slavin & Smith, 2008), small studies in this collection of studies had a much larger effect than larger studies (ES=+0.32 vs. ES=+0.04, respectively). A few possible reasons could explain these findings. First, it is much easier to maintain high implementation fidelity in small-scale studies as compared to large ones. Second, large-scale studies are more likely to use standardized tests, which are often less sensitive to treatment. Lastly, small studies with null effects may have never been written or made available in published or report form, while large-scale studies, especially those funded by the government or non-profit organizations or institutions, are more likely required to make the results, be they positive or negative, available in the public domain as technical reports or in published form.

Limitations

As with any research review, the current review has several limitations. First, only studies with quantitative measures of reading were included. There is much to be learned from other non-experimental studies such as qualitative and correlational research that can add depth and insight to understanding the effects of these technology programs. Second, the review focuses on replicable programs used in realistic school settings over periods of at least 12 weeks, but it does not attend to shorter, more theoretically-driven studies that may also provide useful information, especially to researchers. Third, the review focuses on traditional measures of reading performance, primarily 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

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.

equally likely to be trying to help their students do well on these assessments. However, the review does not report on experimenter-made measures of content taught in the experimental group but not the control group, although results on such measures may also be of importance to researchers or educators. Finally, despite our efforts to locate every qualifying study, only 20 met our standards, making any conclusions tentative.

Implications for Policy and Practice

The most important practical implication of the review presented here is that there is a limited evidence base for the use of technology applications to enhance the reading performance of struggling readers in elementary schools. Only 20 studies met the inclusion standards, and many of these were small experiments; the larger studies, especially those that used random assignment to conditions, reported the smallest effects. Among 8 large, randomized evaluations, the weighted mean effect size was essentially zero (ES=+0.04, n.s.).

Within the existing literature, however, the greatest reason for hope focuses on small-group interventions that supplement first-grade instruction with phonetic activities integrating computer and non-computer activities and occupying substantial time each week. Among currently available models, these include the *Lindamood Phone Sequence Program (LIPS),* and *Read, Write, and Type (RWT).* Among more traditional supplemental CAI models, there is supportive evidence for *Lexia,* also used in first grade. However, each of these is supported by a single, small study, so none can be confidently prescribed as a broadly applicable solution for struggling readers. Further, the effect sizes found for the various technology applications for struggling first graders are, at best, similar to those found for similar phonics-focused small-group interventions that do not use technology, and are much less than those associated with phonetic one-to-one instruction and comprehensive school reform models (see Slavin et al., 2011). For upper-elementary students, none of the technology applications had notable impacts, whereas Slavin et al. (2011) reported substantial positive effects in grades 3-6 for several whole-class interventions, such as cooperative learning, as well as other non-technology interventions targeted on struggling readers.

None of this is meant to imply that technology applications do not have a role in improving outcomes for struggling readers, especially first graders. Three studies that directly compared small-group and individualized tutoring with and without well-integrated technology all found that the use of the technology enhanced reading outcomes for struggling first graders. What it does suggest, however, is that there is no magic in the machine. What determines the effectiveness of technology applications for struggling readers is the nature of the software, the role of the teacher, the nature and quality of professional development and follow-up, the amount of time devoted to the technology and non-technology parts of each approach, how these activities are placed in students' days and weeks, what activities they replace, and much more. If anyone still imagines that computers will make a difference if they merely arrive in a box, ready to plug and play with minimal professional development, the findings reported here should be sobering.

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.

Yet there is no question that technology will be part of future solutions to the problems of reading difficulties. With further research, many of the programs reviewed here could build a stronger evidence base, and the best of them should serve as a basis for further development of impactful models. Approaches using technologies now becoming commonplace in elementary schools, such as interactive whiteboards, electronic response devices, and laptops or other devices for all students, have not yet been adequately researched for struggling readers, but could hold great promise. As computers and other electronic devices become ubiquitous in students' homes, additional possibilities arise in integrating home and school activities. New applications of embedded multimedia, using bits of video to enhance teachers' lessons, also have promise.

The evidence to date shows promise for some types of technology applications, but much more remains to be done both in research and in development of more effective solutions. The problems of reading failure in elementary schools are important, and they justify continued efforts to create and validate reliably effective approaches combining the best efforts of teachers and technology.

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

References marked with an asterisk indicate studies included in the meta-analysis.

Abrami, P. C., & Bernard, R. M. (2006). Research on distance education: In defense of field experiments. *Distance Education 27*(1), 5-26.

*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.

Boardman, A. G., Roberts, G., Vaughn, S., Wexler, J., Murray, C. S., & Kosanovich, M. (2008). Effective instruction for adolescent struggling readers: A practice brief. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). Comprehensive meta-analysis (Version 2). Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-analysis*. West Sussex: Wiley.

Borman, G. D., & Rachuba, L. T. (2009). Evaluation of the Scientific Learning Corporation's Fast ForWord computer-based training program in the Baltimore City Public Schools. *A report prepared for the Abell Foundation*.

Brown, A. & Palincsar, A. (1989). Guided, cooperative learning and individual knowledge acquisition. In Lauren B. Resnick (Ed.), *Knowledge, Learning and Instruction*. Hillsdale, NJ: Lawrence Erlbaum.

*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.

Chambers, B., Slavin, R. E., Madden, N. A., Abrami, P. C., Tucker, B. J., Cheung, A., & Gifford, R. (2008). *The Elementary School Journal, 109* (1), 1-15.

Chambers, B., Slavin, R. E., Madden, N. A., Abrami, P., Logan, M. K., & Gifford, R. (2011). Small-group, computer-assisted tutoring to improve reading outcomes for struggling first and second graders. *The Elementary School Journal, 111* (4), 625-640.

Cheung, A., & Slavin, R. E. (in press). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review*.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445-449.

Clark, R. E. (1985a). Confounding in education computing research. *Journal of Educational Computing Research, 1*(2), 445-460.

Clark, R. E. (1985b). Evidence for confounding in computer-based instruction studies: Analyzing the meta-analyses. *Educational Communication and Technology Journal, 33*(4), 249-262.

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

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

Davidson, J., & Miller, J. (2002). *Scholastic's READ 180: A heritage of research*. New York: Scholastic.

*Drummond, K., Chinen, M., Duncan, T. G., Miller, H. R., Fryer, L., Zmach, C., & Culp, K. (2011). Impact of the *Thinking Reader* Software program on grade 6 reading vocabulary, comprehension, strategies, and motivation. Washington DC: Institute of Education Sciences.

*Dynarski, M., Agodini, R., Heaviside, S. N., Carey, N., Campuzano, L., Means, B., . . . Sussex, W. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort (Vol. aluation)*. Washington, DC: Institute of Education Sciences.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*(1), 262-300.

Hall, L. A. (2004). Comprehending expository text: Promising strategies for struggling readers and students with reading disabilities? *Literacy Research and Instruction, 44*(2), 75-95.

Hall, T. E., Hughes, C. A., & Filbert, C. (2000). C0mputer assisted instruction in reading for students with learning disabilities: A research synthesis. *Education and Treatment of Children, 23*(2), 173-193.

*James-Burdumy, S., Mansfield, W., Deke, J., Carey, N., Lugo-Gil, J., Hershey, A., . . . Faddis, B. (2009). *Effectiveness of selected supplemental reading comprehension interventions: Impacts on a first cohort of fifth-grade students*. (Vol. aluation). Washington, DC: Institute of Education Sciences.

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.

Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children, 70*(3), 299-322.

*Kim, J. S., Capotosto, L., Hartry, A., & Fitzgerald, R. (2011). Can a mixed literacy intervention improve the reading achievement of low-performing elementary school students in an after-school program? Results from a randomized controlled trial of READ 180 Enterprise. *Educational Evaluation and Policy Analysis, 33*(2), 183-201.

*Kim, J. S., Samson, J. F., Fitzgerald, R., & Hartry, A. (2010). A randomized experiment of a mixed-methods literacy intervention for struggling readers in grades 4-6: Effects on word reading efficiency, reading comprehension and vocabulary, and oral reading fluency. *Reading and Writing, 23*(9), 1109-1129. doi: 10.1007/s11145-009-9198-2

Lesnick, J., George, R., Smithgall, C., & Gwynne, J. (2010). *Reading on grade level in third grade: How is it related to high school performance and college enrollment*? Chicago: Chapin Hall at the University of Chicago.

Liao, Y. K. (1999). Effects of hypermedia on students' achievement: A meta-analysis. *Journal of Educational Multimedia and Hypermedia, 8*(3), 255-277.

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

MacArthur, C. A., Ferretti, R. P., Okolo, C. M., & Cavalier, A. R. (2001). Technology applications for students with literacy problems: A critical review. *Elementary School Journal, 101*(3), 273-301.

Macaruso, P., & Hook, P. E. (2001). Auditory processing: Evaluation of FAST FORWORD for children with dyslexia. *IDA Perspectives, 27*(3), 5-8.

*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 Reasearch in Reading, 29*, 162-172.

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

Merzenich, M. M., Jenkins, W. M., Johnson, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing: Evaluation of Fast ForWord for children with dyslexia. *Perspectives, 27*(3), 5-8.

Okolo, C. M., & Bouck, E. C. (2010). Research about assistive technology: 2000-2006. What have we learned? *Journal of Special Education Technology, 22*(3), 19-33.

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.

Pearson, P. D., Ferdig, R. E., Blomeyer, R. L., & Moran, J. (2005). *The effects of technology on reading performance in the middle-school grades: A meta-analysis with recommendations for policy*. Naperville, IL: Learning Point Associates.

*Rabiner, D. L., Murray, D. W., Skinner, A. T., & Malone, P. S. (2010). Randomized trial of two promising computer-based interventions for students with attention difficulties. *Journal of Abnormal Child Psychology, 38*(1), 131-142.

*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.

*Saine, N. L., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2010). Predicting word-level reading fluency outcomes in three contrastive groups: Remedial and computer-assisted remedial reading intervention, and mainstream instruction. *Learning and Individual Differences, 20*(5), 402-414.

Scammacca, N., Roberts, G., S., V., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). Interventions for adolescent struggling readers: A meta-analysis with implications for practice: Portsmouth, NH: RMC Research Corporation, Center on Instruction.

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

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

*Sinkis, D. M. (1993). *A comparison of Chapter One student achievement with and without computer-assisted instruction.* Doctoral thesis, University of Massachusetts.

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

Slavin, R. E. (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.

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research, 79*(4), 1391-1465.

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., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best evidence synthesis. *Educational Research Review, 6*, 1-26.

Slavin, R. E., Lake, C., & Groff, C. (2008). Effective programs in middle and high school mathematics: A best evidence synthesis. *Review of Educational Research, 79*(2), 839-911.

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

Slavin, R. E., & Smith, D. (2008). Effects of sample size on effect size in systematic reviews in education. *Educational Evaluation and Policy Analysis, 31*(4), 500-506.

*Standish, D. (1995). The effects on reading comprehension of Jostens' Integrated Language Arts for second-grade students along with Jostens' Basic Learning System for second-grade Chapter 1 students. *Dissertation Abstracts International, 57*(3), 1079A. (UMI No. 9623238).

Stetter, M. E., & Hughes, M. T. (2010). Computer-assisted instruction to enhance the reading comprehension of struggling readers: A review of the literature. *Journal of Special Education Technology, 25*(4), 1-16.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., . . . Merzenich, M. M. (1996). Language comprehension in languge-learning impaired children improved with acoustically modified speech. *Science, 271*(5), 81-84.

*Torgesen, J., Schirm, A., Castner, L., Vartivarian, S., Mansfield, W., Myers, D., . . . 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: Institute of Education Sciences.

*Torgesen, J., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia, 60*(1), 40-56. doi: 10.1007/s11881-009-0032-y

Vaughn, S., Fletecher, J. M., Francis, D. J., Denton, C. A., Wanzek, J., Wexler, J., . . . Romain, M. A. (2008). Response to intervention with older students with reading difficulties. *Learning and Individual Differences, 18*(3), 338-345.

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing, 23*(8), 889-912.

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.

What Works Clearinghouse. (2012). Beginning reading. What Works Clearinghouse Topic Report. Retrieved July 5, 2012 https://ies.ed.gov/ncee/wwc/

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 1: Educational Technology Applications								
Study	**Design Large/Small**	**Duration**	**N**	**Grade**	**Sample Characteristics**	**Posttest**	**Effect Sizes by Subgroup/ Measure**	**Overall Effect Size**
Supplemental CAI Applications								
Jostens (early form of Compass Learning)								
Sinkis (1993)	Matched (L)	1 year	422 students (228E, 194C)	3, 5, 6	Students in pullout programs in the urban northeast	MAT Comprehension		+0.12
						Grade 3	+0.14	
						Grade 5	+0.22	
						Grade 6	-0.01	
Becker (1994)	Randomized (S)	1 year	1 school 56 students	2-5	Low performing students in inner city Baltimore	CAT		+0.41
Standish (1995)	Matched (S)	1 year	43 students (22E, 21C)	2	Title I students in suburban Delaware	MAT6 Reading		+0.55
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
Captain's Log								
Rabiner et al (2010)	Randomized (S)	1 year	50 students (25E, 25C)	1	Low-SES students with attention difficulties	DIBELS - Fluency	+0.69	+0.40
						Woodcock - Reading	+0.10	
Destination Reading								
Rabiner et al. (2010)	Randomized (S)	1 year	52 students (27E, 25C)	1	Low-SES students with attention difficulties	DIBELS - Fluency	+0.10	+0.12
						Woodcock - Reading	+0.13	
Other Supplemental CAI								
Coomes (1985)	Matched (S)	1 year	4 schools 36 students (18E, 18C)	4	Low achievers in middle class schools in TX. 90% W.	CTBS		+0.30
Bass, Ries, & Sharpe (1986)	Matched (S)	1 year	2 schools (1E, 1C) 145 students (73 E, 72 C)	5-6	Title 1, low SES students from 16 districts in CN, MA, RI.	SRA	+0.22	+0.18
						Virginia Basic Learning Skills Test	+0.13	

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.

Becker (1994)	Randomized (S)	1 year	60 students	2-5	Low achievers in low SES schools in Baltimore, MD; 50% FL	CAT-Reading		+0.10
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 Lowest third of students	SAT-9 Cohort 1 + Cohort 2		-0.07
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 Lowest third of students	SAT-10 Cohort 1 + Cohort 2		+0.04
Comprehensive Models								
READ 180								
Kim et al. (2010)	Randomized (L)	23 weeks	264 students (133E, 131C)	4-6	Struggling readers from three high poverty schools in southeastern MA	TOWRE Word Reading Efficiency GRADE Reading Vocabulary & Comprehension DORF Oral Reading Fluency	+0.03 -0.04 +0.12	+0.03
Kim et al. (2011)	Randomized (L)	23 weeks	306 students (155E, 151C)	4-6	Struggling readers from four high poverty schools in a midsized urban district in southeastern MA	SAT-10 Vocabulary SAT-10 Reading Comprehension DORF Oral Reading Fluency	+0.23 +0.31 +0.10	+0.21
Read About								
James-Burdumy et al. (2009)	Randomized (L)	1 year	871 students (415E, 456C)	5	Low SES students in lowest third of grade	TOSCRF Composite Test Scores GRADE	-0.03 -0.02	-0.03
Small-Group Integrated Supplemental Programs								
Failure Free Reading								
Torgesen et al. (2007)	Randomized (S)	1 year	16 schools 219 students (113E, 104C)	3 and 5	Struggling readers in schools around Pittsburgh; 44%FL, 80%W, 20%AA	Average of Woodcock, TOWRE, AIMSweb, and GRADE 3rd grade 5th grade	 +0.19 -0.05	+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.

Read, Write, and Type-Small Group								
Torgesen et al. (2010)	Randomized (S)	1 year	73 students (34E, 39C)	1	Struggling readers in 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. (2010)	Randomized (S)	1 year	74 students (35E, 39C)	1	Struggling readers in 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	
Thinking Reader								
Drummond et al (2011)	Randomized (L)	1 year	808 students (425E, 383C)	6	Lowest third of students from 16 districts in CN, MA, RI.	GMRT-Vocabulary	+0.14	+0.14
						GMRT-Comprehension	+0.13	
Fast ForWord								
Fast ForWord								
Rouse & Krueger (2004)	Randomized (L)	1 year	4 schools 454 students (237E, 217C)	3-6	Lowest 20% of students in high-poverty schools in Hartford, CT 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	Lowest 25% of students in schools in Appalachian TN 52% FL, 100% W	Terra Nova		+0.15

TABLE 2

Mixed effects moderator analyses examining effect sizes by methodological and substantive features

Descriptors	k	d	95% CI		p	Q-value	df	P
Overall effect size	20	0.14	0.06	0.22	0.00	38.13*	19	0.01
Type of program						3.70	3	0.30
Supplemental	12	0.18	0.04	0.28	0.00			
Comprehensive	3	0.04	-0.09	0.17	0.54			
Small-group	3	0.32	-0.05	0.69	0.09			
Fast For Word	2	0.06	-0.11	0.24	0.48			
Grade level						4.66[a]	2	0.09
Beginning	8	0.36	0.11	0.60	0.00			
Upper Elementary	10	0.07	0.00	0.13	0.04			
Mixed	2	0.25	-0.12	0.61	0.18			
Program intensity						1.20	1	0.27
High (>75min a week)	13	0.19	0.06	0.28	0.00			
Low (<75min a week)	7	0.08	-0.04	0.20	0.19			
Type of publication						6.47*	1	0.01
Published	11	0.25	0.11	0.40	0.00			
Unpublished	9	0.04	-0.03	0.11	0.25			
Year of publication						1.72	3	0.63
1980s	2	0.20	-0.08	0.50	0.17			
1990s	4	0.18	0.01	0.34	0.03			
2000s	7	0.08	-0.05	0.28	0.25			
2010s	7	0.22	0.08	0.36	0.00			
Experimental design						3.58[a]	1	0.06

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.

Randomized	13	0.08	0.01	0.16	0.05			
Quasi-Experiment	7	0.28	0.09	0.48	0.00			
Sample size						11.84*	1	0.001
Small (N<250)	12	0.32	0.17	0.46	0.000			
Large (N≥250)	8	0.04	-0.02	0.10	0.000			
Design & size						11.46*	3	0.00
Large Randomized	7	0.03	-0.03	0.10	0.33			
Small Randomized	6	0.28	0.08	0.48	0.00			
Large Matched Control	1	0.12	-0.07	0.31	0.22			
Small Matched Control	6	0.34	0.13	0.56	0.00			

[a] p<0.10, *p<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.

Appendix
Educational Technology Reading Applications for Struggling: Program Description

Study	Program	Description	Total Mins. Weekly	Intensity
Sinkis (1993) Becker (1994) Standish (1995)	*Jostens*	*Jostens* is an earlier form of an integrated learning system now called Compass Learning. The system is designed to provide an extensive set of assessments, which place students in an individualized instructional sequence. Students work individually on exercises designed to fill in gaps in their skills.	60-150	15-30 minutes per day, 2-5 days per week
Macaruso et al (2006)	*Lexia*	*Lexia Reading* encourages early reading success through the development of critical pre-reading and reading skills. The interactive, phonics-based activities reinforce sound-symbol correspondence and help develop automatic word recognition and comprehension. Activities include the application of reading strategies to single word phrases, sentences, paragraphs, and brief stories. The activity formats also encourage listening skills and following directions. An appealing, deliberately uncluttered interface focuses young students and allows them to easily learn to navigate activities on their own.	40-120	2-4 20-30 minute sessions a week
Rabiner et al (2010)	Two computer-based programs: *Captain's Log* and *Destination Reading*	*Captain's Log* is a commercially available product that provides structured opportunities for exercising attention. *Destination Reading* is a computer-assisted program that targets five key skills: phonemic awareness, phonics, fluency, vocabulary, and comprehension.	100-120	2x50-60-minutes weekly
Coomes (1985)	A variety of software (e.g. *Fundamental Punctuation Practice, Micro-Read, Spelling Program, etc)*	Supplemental CAI in addition to basal reading program. Software provided by several companies coordinated with basal series. Emphasis is on drill and practice.	30	30 minutes weekly

Bass, Ries, & Sharpe (1986)	A variety of software (e.g. *Alpine Skier, Tank Tactics, Big Door Deal*)	Emphasis is on drill and practice.	25	25 minutes weekly
Becker (1994)	*Computer Networking Specialists (CNS)*	*CNS* incorporates a variety of drill & practice and tutorial software from 10+ different independent publishers into its own system for managing assessment and assignment of tasks. The software is based on repetitive practice of isolated skills.	90	3x30 minutes weekly
Dynarski et al (2007) Campuzano et al (2009)	*Destination Reading*	*Destination Reading* includes lessons in phonemic awareness, phonics, vocabulary, fluency, and comprehension. The program is used as a supplement to traditional, teacher-directed instruction, to reinforce reading skills in emergent readers.	25	25 minutes weekly
	Headsprout	*Headsprout® Early Reading* is an Internet-based supplemental early literacy curriculum consisting of eighty 20-minute animated episodes, the first 40 of which are appropriate for prekindergarten age students. The episodes are designed to teach phonemic awareness, phonics, fluency, vocabulary, and comprehension. The program adapts to a child's responses, providing additional instruction and review if a child does not choose the correct answer. Teachers may use stories based on the episodes to reinforce instruction provided in the lessons.	33	33 minutes weekly
	Plato Focus	PLATO Focus is an early reading and listening instruction system that teaches students how to connect sounds with their corresponding symbols in clearly defined sequences. The system features both print materials and instructional technology to teach the five elements of reading: Phonemic awareness, phonics, fluency, vocabulary development, reading comprehension.	No info	No info

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.

	Waterford Early Reading	*Waterford Early Reading Program*™ is a software-based curriculum for students in K-2nd grade. The curriculum is designed to promote reading, writing, and typing, incorporating literacy skills such as letter mastery, language stories, spelling, basic writing skills, reading and listening development, and comprehension strategies. It can be used as a supplement to the regular reading curriculum.	107	107 minutes weekly
	Academy of Reading	*Academy of Reading* is a set of exercises to improve phonemic awareness and sound-symbol association, phonics and decoding skills, fluency and comprehension, and reading proficiency. Students work at their own pace. The software provides assessments for teachers about student usage and progress.	55	55 minutes weekly
	LeapTrack	*LeapTrack* is a supplemental reading product to improve phonemic awareness, phonics, vocabulary, and reading comprehension in addition to other reading skills. Teachers us built-in assessments to identify skills students need to develop and the program provides individualized instruction, including a list of skill cards and books for the student to complete. Students work at their own pace using the *LeapPad, LeapTrack* skills cards, and *LeapFrog*.	30	30 minutes weekly
	Read 180	*READ 180* is a reading intervention program that combines research-based reading practices with technology, providing students a combination of instructional, modeled, and independent reading components. In its full 90-minute version, READ 180 offers a mix of teacher-directed whole group lessons and three 20-minute instructional activities designed to improve word reading efficiency, reading comprehension and vocabulary, and oral reading fluency. A majority of treatment teachers (70%) used *READ 180* as a supplement to their reading curriculum.	100	100 minutes weekly
	KnowledgeBox	*KnowledgeBox* is a server-based collection of resources (text passages, video clips, images, internet sites, software modules) from which teachers can choose resources customized to their local curriculum.	100	100 minutes weekly

Dynarski et al (2007) Campuzano et al (2009)	*Destination Reading*	*Destination Reading* includes lessons in phonemic awareness, phonics, vocabulary, fluency, and comprehension. The program is used as a supplement to traditional, teacher-directed instruction, to reinforce reading skills in emergent readers.	25	25 minutes weekly
	Headsprout	*Headsprout® Early Reading* is an Internet-based supplemental early literacy curriculum consisting of eighty 20-minute animated episodes, the first 40 of which are appropriate for prekindergarten age students. The episodes are designed to teach phonemic awareness, phonics, fluency, vocabulary, and comprehension. The program adapts to a child's responses, providing additional instruction and review if a child does not choose the correct answer. Teachers may use stories based on the episodes to reinforce instruction provided in the lessons.	33	33 minutes weekly
	Plato Focus	PLATO Focus is an early reading and listening instruction system that teaches students how to connect sounds with their corresponding symbols in clearly defined sequences. The system features both print materials and instructional technology to teach the five elements of reading: Phonemic awareness, phonics, fluency, vocabulary development, reading comprehension.	No info	No info
	Waterford Early Reading	*Waterford Early Reading Program*™ is a software-based curriculum for students in K-2nd grade. The curriculum is designed to promote reading, writing, and typing, incorporating literacy skills such as letter mastery, language stories, spelling, basic writing skills, reading and listening development, and comprehension strategies. It can be used as a supplement to the regular reading curriculum.	107	107 minutes weekly

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.

	Academy of Reading	*Academy of Reading* is a set of exercises to improve phonemic awareness and sound-symbol association, phonics and decoding skills, fluency and comprehension, and reading proficiency. Students work at their own pace. The software provides assessments for teachers about student usage and progress.	55	55 minutes weekly
	LeapTrack	*LeapTrack* is a supplemental reading product to improve phonemic awareness, phonics, vocabulary, and reading comprehension in addition to other reading skills. Teachers us built-in assessments to identify skills students need to develop and the program provides individualized instruction, including a list of skill cards and books for the student to complete. Students work at their own pace using the *LeapPad, LeapTrack* skills cards, and *LeapFrog*.	30	30 minutes weekly
	Read 180	*READ 180* is a reading intervention program that combines research-based reading practices with technology, providing students a combination of instructional, modeled, and independent reading components. In its full 90-minute version, READ 180 offers a mix of teacher-directed whole group lessons and three 20-minute instructional activities designed to improve word reading efficiency, reading comprehension and vocabulary, and oral reading fluency. A majority of treatment teachers (70%) used *READ 180* as a supplement to their reading curriculum.	100	100 minutes weekly
	KnowledgeBox	*KnowledgeBox* is a server-based collection of resources (text passages, video clips, images, internet sites, software modules) from which teachers can choose resources customized to their local curriculum.	100	100 minutes weekly

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.

Kim et al (2010) Kim et al (2011)	*Modified version of READ 180*	*READ 180* is a reading intervention program that combines research-based reading practices with technology, providing students a combination of instructional, modeled, and independent reading components. In its full 90-minute version, *READ 180* offers a mix of teacher-directed whole group lessons and three 20-minute instructional activities designed to improve word reading efficiency, reading comprehension and vocabulary, and oral reading fluency. The modified version is a 60-minute after school program, which had all the components in the full version except for teacher-directed whole-group lessons for building vocabulary.	240	4x60 minutes weekly
James-Burdumy et al (2009)	*Read About*	*Read About* is a computer program designed to improve student reading comprehension skills, vocabulary, and content knowledge. In addition, students also use offline materials once per week for 20 minutes. Offline materials include whole-class or small-group lessons on comprehension skills, vocabulary strategies, text types, or writing skills. Students rotate among computer, teacher-led, and independent reading groups. Teacher materials include suggestions for English Language Learners and differentiated instruction.	80	3x20 minutes weekly CAI + 20-minutes weekly non-CAI activities
Torgesen et al (2007)	*Failure Free Reading*	The program combines computer-based lessons, workbook exercises, and teacher-led instruction to teach sight vocabulary, fluency, and comprehension. Students were placed into instructional groups of three students. In addition, the program emphasized building students' vocabulary of sight words rather than phonemic decoding strategies.	165	5x55-minutes weekly
Torgesen et al (2010)	*Read, Write, and Type (RWT) & Lindamond Phoneme Sequence Program (LIPS)*	Both programs were designed to provide explicit and systematic support for the development of phonemic awareness, phonemic decoding, and text reading accuracy in small-group settings. Unlike other CAI programs, these two programs were tightly integrated with classroom instruction.	200	4x50 minutes weekly

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.

Drummond et al (2011)	*Thinking Reader*	*Thinking Reader* is a software program for improving the reading vocabulary and comprehension of students in Grades 5-8. The program embodies an approach to reading instruction known as reciprocal teaching, which requires teachers to model comprehension strategies and support students' efforts to recall and employ those strategies in their reading. *Thinking Reader* is intended to be integrated with classroom discussion and peer interaction.	100-150	5x20-30 minutes weekly
Rouse and Krueger (2004)	*Fast ForWord*	A supplemental reading intervention program designed to develop and strengthen memory, attention, processing rate, and sequencing, by providing immediate positive reinforcement and corrective feedback through graphics.	450-500	5x90-100 minutes weekly
Marion (2004)				

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.