

A Synthesis of Quantitative Research on Reading Programs for Secondary Students

Ariane Baye

University of Liege, Belgium

Amanda Inns

Cynthia Lake

Robert E. Slavin

*Johns Hopkins University, Baltimore,
Maryland, USA*

ABSTRACT

Recent initiatives in the United States and United Kingdom have added greatly to the amount and quality of research on the effectiveness of secondary reading programs, especially programs for struggling readers. In this review of the experimental research on secondary reading programs, the authors focused on 69 studies that used random assignment ($n = 62$) or high-quality quasi-experiments ($n = 7$) to evaluate outcomes of 51 programs on widely accepted measures of reading. Categories of programs using one-to-one and small-group tutoring, cooperative learning, whole-school approaches including organizational reforms such as teacher teams, and writing-focused approaches showed positive outcomes. Individual approaches in a few other categories also showed positive impacts. Categories in which these individual programs appear include programs emphasizing social studies/science, structured strategies, and personalized and group/personalization rotation approaches for struggling readers. Programs that provide a daily extra period of reading and those utilizing technology were no more effective, on average, than programs that did not provide these resources. The findings suggest that secondary readers benefit more from socially and cognitively engaging instruction than from additional reading periods or technology.

The reading performance of students in U.S. middle and high schools is one of the most important problems in education. The National Center for Education Statistics (2016) reported that only 34% of eighth graders scored at or above proficient on the 2015 National Assessment of Educational Progress (NAEP), and at the 12th grade level, 37% of students scored at or above proficient. Secondary reading performance is advancing very slowly. Since 1992, the proportion of U.S. eighth graders scoring at proficient or advanced has risen by only five percentage points and has dropped three percentage points among 12th graders. According to the OECD's (2016) Programme for International Student Assessment, the mean performance of U.S. 15-year-old students was 24th among all countries, and an OECD (2013) survey of adult competencies showed that the average reading level of U.S. young adults (16–24-year-olds) was below the international average for developed countries.

Important gaps continue to exist between groups. Whereas 44% of white eighth graders scored at or above proficient on the NAEP, only 16% of African American students, 21% of Hispanic students, and 22% of American Indian/Alaska Natives did so (National Center for Education Statistics, 2016). Among students qualifying for free or reduced-price lunch, only 20% scored at or above

proficient. At the 12th-grade level, 46% of white 12th graders scored at proficient or better, but the proportion was 17% for African Americans, 25% for Hispanics, and 28% for American Indian/Alaska Natives. In all of these groups, scores have been essentially static since 1992.

The Rising Importance of Evidence

There is a new movement toward the use of evidence in education, especially in the United States and United Kingdom. This movement is intended to identify and disseminate proven programs, thereby improving outcomes for students. The 2015 Every Student Succeeds Act (ESSA) in the United States defines three main levels of evidence, all requiring at least one study with a statistically significant positive effect (and no studies showing negative effects): “Strong” requires at least one randomized study, “moderate” a matched study, and “promising” a correlational study. Programs must meet one of these levels to qualify for some types of federal funding. Educators and policymakers need to know which programs meet ESSA standards, because this evidence is beginning to affect federal, state, and local policies.

The U.S. government has provided substantial funding for development, evaluation, and dissemination of proven programs. The U.S. Department of Education’s Striving Readers program (Boulay, Goodson, Frye, Blocklin, & Price, 2015) focused on secondary reading, and funding from Investing in Innovation (and from the Institute of Education Sciences has also often gone to research on secondary reading (Herrera, Truckenmiller, & Foorman, 2016). The Education Endowment Foundation in England also funds rigorous studies in secondary reading. As a result of these and other investments, the numbers of studies of secondary reading using rigorous research methods, especially cluster-randomized designs with large samples, has increased dramatically.

Current Issues in Secondary Reading

Common Core State Standards

In the United States, secondary reading has been greatly influenced by the Common Core State Standards and other college and career standards. The development and implementation of the Common Core identified explicitly what a literate secondary student should be able to do. The concept of literacy has been broadened to include a wide set of skills (Beach & O’Brien, 2018).

The Common Core has a strong focus on critical literacy: the ability to, for example, evaluate the validity of statements or information in a text, to be open to multiple viewpoints, and to identify and challenge stereotypes (McLaughlin & DeVoogd, 2018). Students should read both literature and informational text and also engage in meaningful conversation and present coherent arguments, expanding literacy beyond written text. The Common Core strongly emphasizes collaboration among students (Guthrie, Klauda, & Ho, 2013). Students must develop sophisticated writing skills in many genres (Graham et al., 2017). The Common Core has introduced reading standards that include disciplinary literacy, taking into account the specificity of disciplinary discourses (Shanahan & Shanahan, 2018). The focus is on how reading and writing differ across content areas.

Technology

The remarkable growth of technology worldwide has also redefined what counts as literacy (U.S. Department of Education, 2017). Students must be conversant not only with traditional written texts but also with the new forms of communication and knowledge sharing that technology offers. Students must be able to navigate texts that may include pictures, games, videos, and blogs and to communicate about them with others (Kervin, Mantei, & Leu, 2018). These multimodal texts are increasingly important to society and are especially relevant to the lives of adolescents, yet are frequently ignored in traditional school settings.

Diversity

In addition to shifts in the texts that students encounter on a daily basis, the profile of adolescents is growing ever more diverse. With the increase in diversity comes an increase in the range of resources that students bring with them to school. The importance of building on the funds of knowledge that diverse students already possess is increasingly recognized as a key feature of instruction.

Civic Engagement

There has also been increased interest in examining how adolescent literacy could support the development of civic engagement (McLaughlin & DeVoogd, 2018). The growth of multimodalities has increased the number of ways that students can engage with their world. In addition, the focus on writing, collaboration, and discussion (as evidenced by the Common Core) provides the tools to increase civic engagement. Specific literacy skills can be developed through action related to issues that students care about (Guthrie, 2015).

The Common Core, a focus on critical literacy, innovative uses of technology, and multimodal text have only begun to show up in research on secondary literacy approaches but often have been studied under different names or as precursors to current practices (see Beach & O'Brien, 2018). Cooperative learning, critical literacy, disciplinary literacy, and multicultural teaching have long histories but have been particularly emphasized in recent research and practice.

The Need for a New Synthesis of Research on Secondary Reading Programs

Over the past decade, several reviews of research on secondary reading programs have provided an important base for the current synthesis. However, the surge in rigorous experiments is so recent that even the most current reviews are not up to date in terms of numbers or methodological quality of studies. As will be seen, in the current review, we found 69 experiments (62 of which were randomized) that met very high evidence standards. In previous reviews of secondary reading, Slavin, Cheung, Groff, & Lake (2008) identified 33 studies that met high methodological standards, and Herrera et al. (2016), using What Works Clearinghouse standards, also identified 33 qualifying studies. Wanzek et al. (2013) found only 10 studies of reading programs for grades 4–12, and Edmonds et al. (2009) located 17. In their review, Flynn, Zheng, and Swanson (2012) reported on only 10 studies of struggling readers in grades 4–8.

The main focus of the current review, therefore, is to learn from the much larger corpus of rigorous evaluations that have become available in recent years. In particular, we used current data to determine which secondary reading approaches would meet ESSA evidence standards, focusing not only on individual programs but also on categories of approaches.

Some previous reviews examined secondary reading interventions and identified approaches associated with particularly positive reading outcomes. These reviews include the one by Slavin et al. (2008), who concluded that secondary reading programs that incorporated cooperative learning and other innovations in classroom teaching practices had the strongest effects on reading achievement in grades 6–12. Herrera et al. (2016) also reported that cooperative learning approaches and other methods providing extensive professional development (PD) were particularly likely to have positive outcomes. Dietrichson, Bøg, Filges, and Jørgensen (2017) found the strongest support for tutoring, feedback/progress monitoring, and cooperative learning.

Focus of the Review

The present review synthesizes research on reading outcomes of programs designed for middle and high school students. We used best evidence synthesis (Slavin, 1986), a method adapted from meta-analysis (see Cooper, 1998; Lipsey & Wilson, 2001) that includes narrative and numeric summaries of the methods and findings of all studies meeting a common set of inclusion criteria. In the review, we describe methods and outcomes of individual studies and programs and place studies in well-justified categories to find patterns that may have broader applicability and may suggest where additional development and research may be most fruitful.

Limitations

It is important to note that the focus of this review is squarely on rigorous, mostly randomized, quantitative research evaluating secondary reading programs. There has been much other research on secondary reading, including qualitative and correlational methods and outcomes other than standardized tests. These studies are important in building theory and understanding. However, in light of recent developments in evidence-based reform, it is crucial to have a clear understanding of which programs and program types are able to accelerate the reading achievement of secondary students, in studies that would meet congressionally defined standards per ESSA, and that is what we have attempted to provide.

Method

Criteria for Inclusion

We focused on a set of studies that met rigorous inclusion criteria, derived from the ESSA evidence standards for ratings of strong and moderate. We designed 12 criteria to minimize bias and maximize potential replicability in schools not involved in the research:

1. Studies evaluated reading programs for middle and high schools, grades 6–12.
2. Students who qualified for special education services but attended mainstream English or reading classes were included.
3. Studies compared students in a given reading program with those taught in an alternative or business-as-usual control group.
4. Studies could have taken place in any country, but the report had to be available in English. In practice, all included programs took place in the United States or United Kingdom.

5. Studies used random assignment to experimental and control conditions, or quasi-experimental methods in which treatment assignments were specified in advance.
6. Studies had to provide pretest data. Those with experimental-control differences equivalent to an effect size of 0.25 or more on pretests were excluded. Pretest equivalence had to be acceptable both initially and based on pretests for the final sample, after attrition. Differential attrition from pre- to posttest had to be less than 15%.
7. Treatments had to be delivered by ordinary teachers, not by researchers, because effect sizes are inflated when researchers deliver the treatment (Scammacca et al., 2007).
8. Studies' dependent measures had to be quantitative measures of reading performance. When standardized tests were used, total reading or total comprehension scores were accepted. If comprehension and vocabulary measures were presented separately, a total reading score was computed weighting comprehension at twice the value of other measures to correspond to weightings of these factors in most standardized tests for the secondary grades.
9. Assessments made by developers or researchers were excluded, as such measures have been found to greatly overstate program impacts

(Cheung & Slavin, 2016; de Boer, Donker, & van der Werf, 2014; Edmonds et al., 2009).

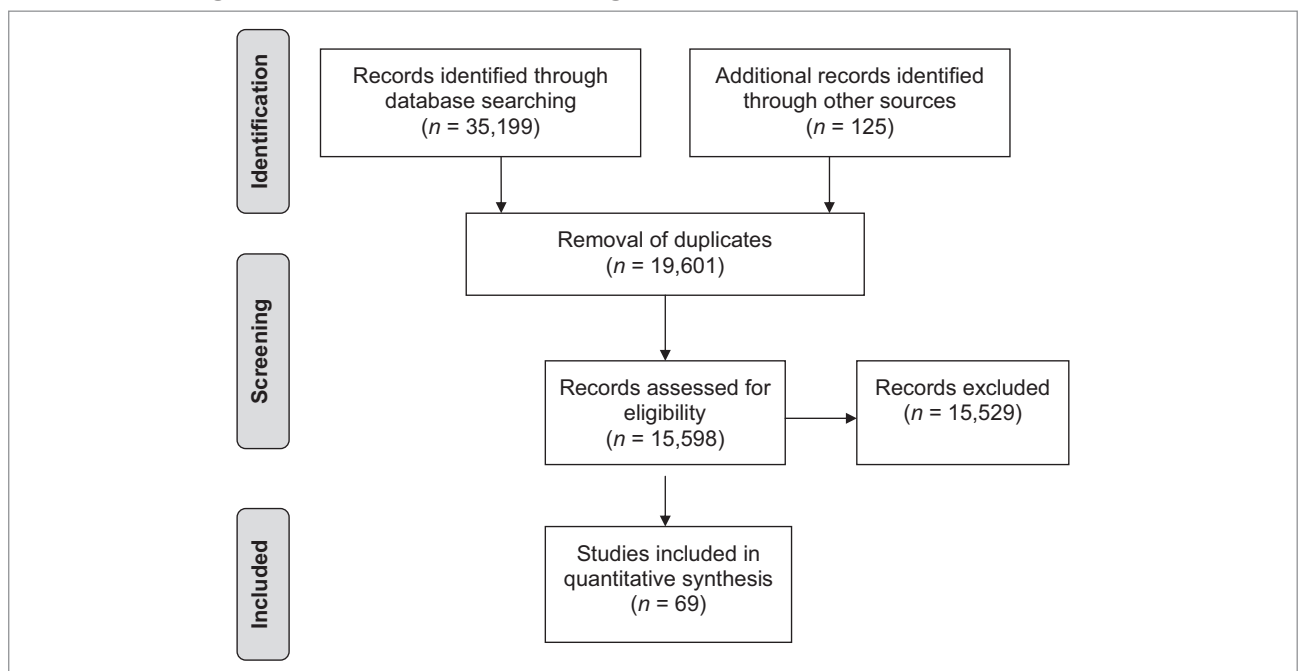
10. Studies had to have a minimum duration of 12 weeks to make it more likely that effective programs could be replicated over extended periods.
11. Studies had to have at least two teachers and 30 students in each treatment group.
12. Studies had to be carried out after 1990, but for technology approaches, we used a start date of 2000 because of the significant advances in technology since then.

Literature Search and Selection Procedures

We conducted a broad literature search in an attempt to locate every study that could possibly meet the inclusion requirements, and then we screened studies to determine whether they were eligible for review. The process is summarized in Figure 1.

We carried out the comprehensive search of the literature through a multistep process: (1) an electronic database search, (2) a hand search of 10 peer-reviewed journals, (3) an ancestral search of recent meta-analyses, (4) a web-based search of educational research websites and educational publishers' websites, and (5) a final review of citations found in relevant documents retrieved from the first search wave.

FIGURE 1
Flowchart Showing the Number of Studies at Each Stage of the Search and Review Process



Note. This process was conducted per Moher, Liberati, Tetzlaff, and Altman (2009).

First, we conducted electronic searches of educational databases (i.e., ERIC, Education Full Text, PsycINFO, ProQuest Education Database, ProQuest Dissertations & Theses Global, Google Scholar) to locate studies published between 1990 and 2017. We used key terms related to secondary reading (i.e., *literacy, reading, middle school, high school*) in various combinations with intervention- and evaluation-related terms (i.e., *program effectiveness, randomized controlled trial, control group, impact, student achievement, quasi-experiment*) in an effort to capture the comprehensive body of research on reading programs for secondary students. A complete list of key words is shown in Table 1.

Following the electronic search, we conducted hand searches of recent tables of contents of 10 key reading and educational research journals from 2013 to 2017: *Reading Research Quarterly, Journal of Literacy Research, Scientific Studies of Reading, Journal of Adolescent & Adult Literacy, Reading & Writing Quarterly, Review of Educational Research, Journal of Educational Psychology, The Journal of Educational Research, American Educational Research Journal, and Journal of Research on Educational Effectiveness*. We cross-referenced relevant citations with the results of the literature search in an effort to corroborate that the search terms were returning the pertinent studies. In addition, we also cross-referenced citations from recent meta-analyses (Edmonds et al., 2009; Herrera et al., 2016; Paul & Clarke, 2016; Scammacca et al., 2016; Wexler, Vaughn, Edmonds, & Reutebuch, 2008) with the search results and reviewed any additional studies.

We also examined web-based repositories and educational publishers' websites (e.g., MacMillan/McGraw-Hill, Pearson, Houghton Mifflin Harcourt) to identify unpublished studies because of the known difference in effect sizes between published and unpublished studies (Polanin, Tanner-Smith, & Hennessy,

2016). We searched for studies reviewed by the What Works Clearinghouse and ones reported online by Investing in Innovation, Institute of Education Sciences, Education Endowment Foundation, and other funders and researchers. We contacted producers and developers of reading programs to check whether they knew of studies that we had missed. Finally, we reviewed citations of all documents found to be particularly relevant to the review for any other possible studies. The process continued until no new studies were found.

We first screened each study by examining the title and abstract to identify whether the study met the inclusion criteria. For studies that could not be immediately eliminated during this process, one of us read the full text. We retained studies that met the inclusion criteria and any others where this was not clear. A different author then examined the retained studies to confirm that they met all inclusion criteria. Any disagreements were discussed by the previous two authors and a third author, and decisions were made once consensus was reached. Sometimes key information was missing, and we asked study authors to provide additional data.

After removing duplicate listings, these search strategies yielded 15,598 studies for screening. A review of the abstracts and full-text articles, when necessary, eliminated 15,529 studies, which left 69 articles that met inclusion standards per the focus of this review. Figure 1 shows the process by which articles were identified and screened to result in the final set.

Coding

We coded studies that were determined to have met the inclusion criteria for several features: program components, effect sizes, demographic and methodological

TABLE 1
Search Terms Used in Database Searches

Sample description	Intervention description	Study design description
<ul style="list-style-type: none"> • <i>Adolescent</i> • <i>Elementary secondary school</i> • <i>High school</i> • <i>Middle school</i> • <i>Secondary school</i> • <i>Student achievement</i> 	<ul style="list-style-type: none"> • <i>Language arts</i> • <i>Literacy</i> • <i>Reading</i> • <i>Reading instruction</i> • <i>Reading strategies</i> 	<ul style="list-style-type: none"> • <i>Comparative analysis</i> • <i>Control group</i> • <i>Fidelity</i> • <i>Impact</i> • <i>Instructional effectiveness</i> • <i>Intervention</i> • <i>Program effectiveness</i> • <i>Program evaluation</i> • <i>Quasi-experiment</i> • <i>Randomized controlled trial</i> • <i>Randomized evaluation</i> • <i>Randomized experiment</i>

Note. The terms were used in various combinations.

features, and cross-cutting factors. Each of these areas is discussed in detail later. Studies were coded by one author, and then codes were verified by another author. Disagreements were decided in discussion with a third author until consensus was reached.

Effect Sizes

We computed effect sizes as the difference between individual experimental and control student posttests after adjustment for pretests and other covariates, divided by the unadjusted posttest control group standard deviation. If the control group standard deviation was not available, we used a pooled standard deviation. We used procedures described by Lipsey and Wilson (2001) to estimate effect sizes when unadjusted standard deviations were not available. If pretest and posttest means and standard deviations were presented but adjusted means were not, we subtracted effect sizes for pretests from effect sizes for posttests.

Statistical Significance

Statistical significance is reported for all studies. The criteria for statistical significance are generally those of the What Works Clearinghouse (2017). When studies used random assignment or matched assignment at the individual level, they usually compared experimental and control groups using analysis of covariance, controlling for pretests and, in most studies, demographic variables (e.g., race, free or reduced-price lunch).

When studies randomly assigned classes or schools to treatments or compared matched classes or schools, they should have used multilevel modeling such as hierarchical linear modeling (Raudenbush & Bryk, 2002) to analyze the data. However, if a clustered design mistakenly used a student-level analysis, the review recalculated the analysis to estimate the results that would have been obtained in hierarchical linear modeling, using a formula provided by the What Works Clearinghouse (2017). Following ESSA evidence standards, a program is considered effective if it has at least one statistically significant positive effect and no statistically significant negative effects.

Statistical Procedures

We calculated mean effect sizes across studies after assigning each study a weight based on inverse variance (Lipsey & Wilson, 2001), adjusted as suggested by Hedges (2007). In combining across studies, we used a random effects model as recommended by Borenstein, Hedges, Higgins, and Rothstein (2009) when there was a belief that there was no single true effect size but a range of effect sizes that may have depended on other factors. We calculated weighted mean effect sizes and

meta-analytic tests such as *Q* statistics in R (R Core Team, 2016) using the *metafor* package (Viechtbauer, 2010).

Program Categories

Once inclusion criteria were met, we placed each program's evaluated studies in categories according to the most important and distinctive components, the key elements of a program that distinguish it from other programs. Category assignments were based on independent close reading of articles and websites by us and then debated and resolved by consensus. All of us read all accepted studies and reread and debated if there were any differences about category assignments. We also consulted program websites, nonresearch writings about programs, and research citations to each program to confirm category assignments.

Research and Theory Supporting Main Program Components

The identified program components and resulting categories were guided by two main sources. The first, Reading Next (Biancarosa & Snow, 2006), identified 15 elements of adolescent literacy programs that support increased achievement. The second was the Institute of Education Sciences practice guide on adolescent literacy (Kamil et al., 2008). We combined and adapted the recommendations of these reports to produce 10 categories of programs, each with distinguishing components supported by prior research and theory.

The Importance of Metacognitive Strategies

The dominant theory in the teaching of secondary reading has for many years focused on the teaching of metacognitive strategies. This approach is intended to help students become aware of their own thinking and to use specific heuristics (e.g., clarification, summarization, graphic organizers, outlining, prediction) to help students comprehend what they read. Students are taught which strategies to use for comprehending different types of text. Reading comprehension requires readers to integrate text with what they already know to derive meaning, which requires flexible use of multiple strategies. However, this process is for the most part invisible, so teachers must explicitly teach students how to use these strategies to make sense of text (e.g., Biancarosa & Snow, 2006; Dignath & Büttner, 2008; Frankel, Becker, Rowe, & Pearson, 2016; Kamil et al., 2008; Pressley, 2003). Reading strategies may be different for different disciplines and reading genres, and students need to know when to use each (Shanahan & Shanahan, 2018).

All qualifying programs in this review incorporated metacognitive strategies to a significant degree. For this reason, we assumed a focus on metacognitive strategies of one kind or another for all programs in this review, because there are few, if any, alternative programs to use as a point of comparison.

The Importance of PD

All of the accepted studies provided significant PD to teachers. PD is an essential element of school improvement (see Desimone & Garet, 2015). Some studies explained the amount and quality of PD provided to teachers, but most did not, so we were unable to use the amount of PD as a cross-cutting factor.

The 10 Program Categories

Note that the number of programs and studies in each category vary greatly, making conclusions about those with the fewest studies less conclusive than those with larger numbers.

Tutoring

In tutoring programs, struggling readers receive one-to-one or small-group tutoring in groups of one to four students. Tutors can be teachers, paraprofessionals, volunteers, or older students. Tutoring sessions are typically given either on some proportion of days (as few as once a week) or daily for a few months. One-to-one and small-group tutoring by teachers and paraprofessionals have been very effective in the elementary grades (Edmonds et al., 2009; Inns, Lake, Pellegrini, & Slavin, 2018; Slavin, Lake, Davis, & Madden, 2011; Wanzek & Vaughn, 2007), with more limited evidence supporting its effectiveness in secondary school (Jun, Ramirez, & Cumming, 2010; Wanzek et al., 2013).

Tutoring emphasizes personalization to the needs and learning strengths of individual students, opportunities to vary the level and pace of instruction for students, and forming personal, caring relationships between tutors and students.

Cooperative Learning Programs

These programs involve students working daily in small mixed-ability groups, with usually four or five members each. The students are encouraged to help one another learn academic content, especially helping one another learn and apply metacognitive comprehension strategies. Cooperative learning approaches have been shown to be among the most effective strategies for improving adolescent literacy (Dietrichson et al., 2017; Herrera et al., 2016; Slavin et al., 2008).

Cooperative learning may improve reading comprehension in several ways. First, it emphasizes motivation through engagement with peers and encouragement

from them, learning by explaining to peers and receiving explanations from them, and personalization through individualized feedback from peers and teachers (Roseth, Johnson, & Johnson, 2008; Slavin, 2015). Cooperative learning can also provide opportunities for participation in high-quality discussions of text that support and increase comprehension (Guthrie, 2015; Kamil et al., 2008) and corresponds with the emphasis by Frankel et al. (2016) on motivation and engagement.

Whole-School Approaches

Several programs provide PD for teachers of all major subjects, in a coordinated schoolwide plan to increase student motivation and achievement. These programs may also build cross-disciplinary teaching teams and focus on socioemotional skills and motivation (Guthrie, 2015; Guthrie et al., 2013). Previous research on comprehensive school reforms has demonstrated that some of these programs are effective when well implemented (G.D. Borman, Hewes, Overman, & Brown, 2003). In secondary reading, providing all teachers with PD in reading comprehension strategies is expected to give teachers across disciplines a common language and tool kit of effective strategies, giving students consistent comprehension instruction in many of their classes.

Writing-Focused Approaches

Writing is a critical skill in itself, of course, but there is evidence that a focus on writing can also increase reading comprehension (see, e.g., Graham et al., 2017; Graham & Hebert, 2011). Writing about text may help improve comprehension (Fitzgerald & Shanahan, 2000). Classroom activities focused on writing instruction, such as specific writing skills instruction or learning about the writing process, may support the development of related reading skills. Both reading and writing are communication processes, so in learning how to write, students understand the communication process and may become better at all aspects of communication (Hodges, Feng, Kuo, & McTigue, 2016). Further, an emphasis on writing engages secondary learners in self-expression, making learning literacy more active and social than learning only from reading (Graham et al., 2017).

Content-Focused Approaches

Shanahan and Shanahan (2018) emphasized that reading is disciplinary. In secondary schools, metacognitive strategies may be tailored to informational text, especially in social studies or science. This reading instruction may be provided by English, reading, social studies, or science teachers.

Integrating reading within content areas can improve literacy outcomes (Biancarosa & Snow, 2006;

Langer, 2001). The rationale is to focus teaching on the specific requirements of the texts so students can soon apply their new skills to their content classes, facilitating transfer and retention. Focusing on specific types of text is important because disciplinary texts have specialized demands and require tailored comprehension skills (Lee & Spratley, 2010). This leads to students learning how to, for example, read like a scientist or historian, depending on the requirements of the text (Goldman, 2012; Shanahan & Shanahan, 2018). A recent review reported positive effects of teaching expository text structure on comprehension skills in elementary and secondary schools (Pyle et al., 2017).

Vocabulary-Focused Approaches

Several programs focus on developing students' vocabulary, expecting that it will generalize to building students' comprehension. These programs explicitly teach selected vocabulary to students and include strategies to derive the meaning of unknown words, such as through context clues. Vocabulary proficiency has been shown to have a relation with reading proficiency (Oslund, Clemens, Simmons, & Simmons, 2018; Uccelli, Phillips Galloway, Barr, Meneses, & Dobbs, 2015). In their review of research on vocabulary interventions in grades K–12, Wright and Cervetti (2017) found positive effects on comprehension of passages containing specifically taught words but not on generalized comprehension.

Strategy-Focused Instruction

The Strategic Instruction Model (SIM) is a family of programs that emphasize teaching students step-by-step approaches to cope with comprehension, decoding, writing, and other difficulties. These metacognitive strategies include summarizing, questioning, identifying the main idea, and using graphic organizers. Students learn mnemonics to recall how to accomplish key metacognitive objectives (Deshler & Schumaker, 2005). As noted earlier, teaching students these comprehension strategies can have a positive impact on reading achievement (Pearson & Cervetti, 2013; Shanahan & Shanahan, 2018).

Personalization Approaches

Personalization refers to instructional approaches that adapt to the learning levels, interests, or other characteristics of individual students. Historically, this concept was captured by the term *individualization*, where students are placed at a curricular point appropriate to their individual achievement level and then proceed at their own rate through the material, with frequent checks for understanding. The term *personalization* adds to this additional factors to which instruction may be adapted, such as interests, learning preferences, and ideal learning modes (U.S. Department of Education, 2017). For example, Accelerated Reader and iLit provide

students with choices among hundreds of books at their own instructional level, so the software is personalized both for reading level and for student interests.

Today, personalized approaches usually use computers, but this is not a defining characteristic. For example, Accelerated Reader existed long before e-books were available.

The rationale for personalization is primarily that students will learn better and faster if the material they read and respond to is at their learning level, within their zone of proximal development (Vygotsky, 1978). In addition, personalized content may avoid frustrating students with work that they cannot do or boring them with work that is too easy. Content that contains an element of choice is likely to be more motivating and interesting (Guthrie, 2015; Stipek, 2002).

Group/Personalization Rotation Approaches

Several programs rotate students through activities. Typically, one activity is teacher directed (e.g., a traditional lesson), and one is personalized (usually computer-assisted instruction). A third activity may involve cooperative learning. The idea is to use each setting to accomplish goals for which it is ideal. For example, teachers may be best at explaining new or difficult ideas, computers may be ideal for providing personalized practice, and cooperative learning may be ideal for peer tutoring or projects (U.S. Department of Education, 2017).

Intensive Group Approaches

Although most secondary reading approaches can be used with struggling readers, there is a particular category of approaches uniquely focused on the needs of students lacking key skills that should have been learned in elementary school, especially decoding and fluency. Such programs also focus on comprehension, vocabulary, and other objectives but are distinctively focused on identifying and remediating serious gaps in students' prior learning (Vaughn & Fletcher, 2012).

Cross-Cutting Factors

In addition to the 10 strategy types that we just discussed, we also analyzed two important cross-cutting factors: extra daily periods for reading instruction and use of technology.

Extra Class Periods for Reading Instruction

Many of the qualifying programs were provided to students in daily class periods in addition to ordinary reading or English periods, replacing music, art, study hall, or other ordinary parts of the school day. Teachers of the additional period were invariably certified. Such double-dose strategies for reading have been popular as

a response to accountability pressure to improve outcomes. Evidence on extra-time programs is mixed. Studies of double-dose programs in ninth-grade English (Nomi, 2015) found positive impacts on multiple outcomes, including course grades, standardized tests, and graduation rates, with larger impacts for initially lower performing students. One study in middle school (Dougherty, 2013) found that although there were immediate benefits of double-dose reading in sixth grade, the impacts disappeared by eighth grade. However, in reviews of studies of extra-time programs, Kidron and Lindsay (2014) and Zief, Lauver, and Maynard (2006) failed to find positive effects.

Programs Incorporating Technology

A number of widely used secondary reading programs incorporate technology. These programs vary greatly, so rationales depend on what technology is intended to accomplish in a given approach. Most provide self-paced instruction at students' reading levels, with immediate feedback and rewards as students progress. The rationales for such approaches are essentially those given previously for personalization approaches. Some of these programs, such as READ 180 and Passport Reading Journeys, integrate whole-group and small-group instruction with computer-based practice and instruction. Other programs use technology as a means to deliver instruction and practice that is additional to the normal classroom curriculum, such as the e-books and activities included with Accelerated Reader and iLit. Technology may also be used as a tool across all or most subjects to facilitate higher order skills, as in eMINTS. Some types of educational technology have been shown to be effective for secondary students in literacy (Cheung & Slavin, 2013, 2016). We distributed programs using technology among categories according to what they do, rather than in a separate technology category, but we also computed mean effect sizes across categories for all technology approaches.

Demographic and Methodological Moderators

Four important moderators were identified, which had sufficient variation within and/or across studies to enable statistical comparisons.

Struggling Readers

Almost all of the qualifying studies took place in high-poverty schools serving many struggling readers. Studies of tutoring focused on struggling readers. In addition, some studies made a direct comparison between outcomes for struggling readers and outcomes for all students. Thirty-six studies allowed for such a comparison.

Approaches for struggling readers varied widely. Beyond tutoring, the most common approach was to provide struggling readers with an extra daily period of reading instruction, as described previously. Programs using personalization approaches (mostly computer-assisted instruction) sought to assist struggling readers by providing them with individualized content that they could complete at their own pace. Other approaches had components likely to be especially beneficial to struggling readers but were embedded in teaching to whole classes. For example, cooperative learning and writing-focused methods put students in mixed-ability groups in which struggling readers receive assistance and modeling from peers. Strategy models, such as SIM and content-focused approaches, teach all students specific strategies for comprehension, but these may be more important for readers who would not otherwise discover these strategies on their own (see Frankel, Jaeger, Brooks, & Randel, 2015).

Middle/High School

We compared studies that took place in grades 6–8 with those in grades 9–12. It is important to note that most high school studies actually involved grades 9 and 10, with grades 11 and 12 studied less often.

Randomized/Quasi-experimental Designs

Most studies used randomized designs, but we wanted to know whether there were different outcomes for quasi-experiments (matched studies).

Clustered/Student Level

Clustered designs have become far more common in recent years. We wanted to know whether such studies differed in mean effect sizes compared with studies with assignment at the student level (excluding tutoring studies, which were always at the student level).

Results

A total of 69 studies evaluating 51 different programs met the criteria of this review (note that when two distinct programs were compared with control groups and reported in the same article, we counted them as two studies). As a group, the studies were of very high methodological quality. Sixty-two (90%) used random assignment, and only seven (10%) used matched, quasi-experimental designs. In 34 studies (49%), the unit of analysis was the school or classroom, and in each case, analyses were appropriate to the level of clustering (or corrected to be so). Across all studies, the weighted mean effect size was 0.09 ($p < .001$). Nineteen studies found statistically significant positive effects, and across these, the average effect size

was 0.19. Because the NAEP gap between African American and white students and between Hispanic and white students is about a half standard deviation, using a program with an effect size of 0.19 could reduce the achievement gap by about 38%. Table 2 summarizes effect sizes and other information for the 10 categories, the two cross-cutting factors, and the four moderators.

The Appendix describes individual programs and studies. Tables 3–12 present characteristics and outcomes of all accepted studies, grouped by categories.

Tutoring Interventions

One-to-one or small-group tutoring is a widely used and effective intervention for struggling readers in

elementary schools (Inns et al., 2018; Slavin et al., 2011) but is rarely used in secondary schools. All seven of the qualifying tutoring studies (see Table 3) were done in England, as a result of a funding initiative focused on helping struggling students enter secondary school with adequate reading skills.

Taken together, the overall weighted effect size for tutoring programs provided by teachers or paraprofessionals was 0.24 ($p < .001$). Effect sizes were 0.28 ($p < .001$) for three one-to-one programs and 0.14 (nonsignificant [n.s.]) for two small-group programs. One of the small-group studies, however, took place over a period of only 12 weeks and focused on phonics, whereas the final test focused on comprehension (King & Kasim, 2015). A single study of

TABLE 2
Summary of Effect Sizes by Program Category

Program	<i>k</i>	Effect size	95% confidence interval	<i>Q</i>	<i>I</i> ²	χ^2
Tutoring by paid adults	5	0.24***	[0.10, 0.38]	7.28	41.99	0.01
• Tutoring: One-to-one	3	0.28***	[0.12, 0.45]	3.19	39.85	0.01
• Tutoring: Small-group	2	0.14	[-0.20, 0.48]	3.44	70.96	0.04
Cooperative learning	6	0.10*	[0.00, 0.20]	1.74	0.00	0.00
Whole school	8	0.06*	[0.00, 0.13]	3.22	0.00	0.00
• With organizational elements	3	0.09*	[0.02, 0.17]	0.73	0.00	0.00
• Without organizational elements	5	0.00	[-0.12, 0.11]	0.64	0.00	0.00
Writing focused	3	0.13***	[0.08, 0.17]	0.64	0.00	0.00
Content focused	7	0.08*	[0.01, 0.15]	3.32	0.00	0.00
Vocabulary focused	4	0.06	[-0.06, 0.18]	0.20	0.00	0.00
Strategy-focused instruction	8	0.09***	[0.04, 0.14]	4.06	0.00	0.00
Personalization	6	0.13*	[0.02, 0.23]	3.49	0.00	0.00
Group/personalization rotation	17	0.09***	[0.04, 0.13]	9.85	0.00	0.00
Intensive group approaches	3	0.00	[-0.10, 0.10]	1.49	0.00	0.00
<i>Cross-cutting factors</i>						
Additional reading period (yes = 27, no = 42)	69	-0.01	[-0.05, 0.03]	52.69	0.00	0.00
Technology application (yes = 23, no = 46)	69	-0.01	[-0.06, 0.04]	52.52	0.00	0.00
<i>Demographic and methodological moderators</i>						
Struggling readers (<i>n</i> = 36)/all students (<i>n</i> = 36)	72	0.01	[-0.03, 0.05]	56.58	2.54	0.00
Middle (<i>n</i> = 44)/high school (<i>n</i> = 30)	74	0.00	[-0.04, 0.05]	57.74	0.00	0.00
Randomized (<i>n</i> = 62)/quasi-experimental designs (<i>n</i> = 7)	69	-0.04	[-0.10, 0.01]	50.08	0.00	0.00
Clustered (<i>n</i> = 34)/student-level (<i>n</i> = 35)	69	-0.05	[-0.10, 0.00]	49.30	0.00	0.00

* $p < .05$. *** $p < .001$.

TABLE 3
Tutoring Interventions for Struggling Readers

Intervention	Design/ treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
One-to-one tutoring								
<i>Catch Up Literacy</i>								
Rutt et al. (2015)	SR	30 weeks	Students: 557	Years 6 and 7	United Kingdom; 21% FRL	NGRT	0.16	0.16
<i>Perry Beeches Coaching Programme</i>								
Lord et al. (2015)	SR	1 year	Students: 291	Year 7	United Kingdom; 55% W; 20% EL; 58% FRL	Progress Test in English (GL Assessment)	0.36*	0.36*
<i>REACH (tutoring)</i>								
Sibieta (2016)	SR	20 weeks	Students: 202	Years 7 and 8	United Kingdom; 68% W; 63% SPED; 24% EL; 31% FRL	NGRT	0.42*	0.42*
Small-group tutoring								
<i>Butterfly Phonics</i>								
Merrell & Kasim (2015)	SR	4 months	Students: 310	Year 7	United Kingdom; 78% W, 16% AA; 35% SPED; 64% EL; 51% FRL	NGRT	0.30*	0.30*
<i>Rapid Phonics combined with Sound Discovery</i>								
King & Kasim (2015)	SR	12 weeks	Students: 178	Years 6 and 7	United Kingdom; 50% W; 50% SPED; 50% EL; 50% FRL	NGRT	-0.05	-0.05
Volunteer tutoring								
<i>TextNow Transition Programme</i>								
Maxwell et al. (2014)	SR	15 weeks	Students: 391	Years 6 and 7	United Kingdom; 16% EL; 25% FRL	NGRT	-0.06	-0.06
Cross-age tutoring								
<i>Paired Reading</i>								
Lloyd et al. (2015)	CR	16 weeks	Classes: 120 Students: 1,306	Year 7	United Kingdom; 96% W; 20% FR	NGRT	-0.02	-0.02

Note. Design/treatment: CR = cluster randomized; SR = student randomized. Demographics: AA = African American; EL = English learner; FRL = receiving free or reduced-price lunch; SPED = special education; W = white. Measure: NGRT = New Group Reading Test (United Kingdom). * $p < .05$ at the appropriate level of analysis (cluster or individual).

volunteer tutoring and one of cross-age tutoring found no statistically significant effects on reading.

As an example of one-to-one tutoring, Catch Up Literacy is a structured one-to-one tutoring intervention. Paraprofessionals provide 15-minute sessions to struggling readers twice a week over the course of a school year. Each session includes prepared reading, reading aloud, discussing the text, and linked writing. A study of Catch Up Literacy in year 6 (Rutt, Kettlewell, &

Bernardenelli, 2015) found a marginally statistically significant difference favoring the tutored students (effect size [ES] = 0.16, $p = .08$).

An example of one-to-small group tutoring is Butterfly Phonics, which uses formal phonics instruction, teaching the global aspects of text, and class discussion of text meaning, to improve reading comprehension. The program is delivered to groups of six to eight students by a trained practitioner and an assistant. Merrell and

Kasim (2015) evaluated Butterfly Phonics and found an effect size of 0.30 ($p < .001$).

Cooperative Learning Approaches

Cooperative learning methods, in which students work in small groups to help one another grow in reading skills, are widely used in reading and in many other subjects in all grades (Slavin, 2015). Six studies of two cooperative learning programs are shown in Table 4. The weighted mean effect size for all cooperative learning studies was 0.10 ($p < .05$).

The Reading Edge is an example of cooperative learning in secondary reading. In the program, middle school students work in four- or five-member teams to help one another build reading skills. Students engage in partner reading, story retelling, story-related writing, word mastery, and story structure activities to prepare themselves and their teammates for individual assessments that form the basis for team scores. Instruction focuses on explicit teaching of metacognitive strategies.

Across three studies of The Reading Edge, the weighted mean effect size was 0.15 (Slavin, Chamberlain, Daniels, & Madden, 2009; Stevens & Durkin, 1992, studies 1 and 2). The outcomes in Slavin et al.’s study were statistically significantly positive ($ES = 0.15, p < .05$).

In contrast, three studies of Collaborative Strategic Reading found a mean effect size of 0.04. The difference in outcomes between The Reading Edge and Collaborative Strategic Reading may have to do with the cooperative learning structures used in each model. The Reading Edge has a strong emphasis on teams, with group goals based on gains made by all individual team members. Collaborative Strategic Reading uses groups but not team goals. Studies have found that team goals are crucial to the success of cooperative learning (Slavin, 2015).

Whole-School Approaches

Whole-school approaches provide PD to teachers across entire schools or grade levels. Two of these approaches emphasize generic teaching methods, school organization

TABLE 4
Cooperative Learning Approaches

Intervention	Design/treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>The Reading Edge/student team reading</i>								
Slavin et al. (2009)	SR	1 year	Students: 788	6	90% W; 15% SPED; 61% FRL	GMRT	0.15*	0.15*
Stevens & Durkin (1992) study 1	CQE	1 year	Schools: 5 Students: 3,986	6–8	High poverty; majority AA	CAT		0.38
Stevens & Durkin (1992) study 2	CQE	1 year	Schools: 6 Classes: 59 Students: 1,223	6	75% AA; 58% FRL	CAT		0.08
<i>Collaborative Strategic Reading</i>								
Denver Public Schools (2016)	CR	1 year	Schools : 16 Students : 5,660	6–8	62% H, 19% W, 11% AA; 30% EL; 11% SPED; 76% FRL	GMRT Transitional Colorado Assessment Program	0.03 0.02	0.03
Vaughn et al. (2011)	CR	18 weeks	Classes: 61 Students: 782	7 and 8	51% H, 43% W; 52% FRL	GMRT AIMSweb maze TOSREC	0.12 –0.08 0.07	0.04
Vaughn et al. (2013)	CR	20 weeks	Classes: 48 Students: 472	7 and 8	51% W, 42% H; 6% LEP; 7% SPED	GMRT TOSREC	0.10 0.11	0.10

Note. Design/treatment: CQE = cluster quasi-experimental design; CR = cluster randomized; SR = student randomized. Demographics: AA = African American; EL = English learner; FRL = receiving free or reduced-price lunch; H = Hispanic; LEP = limited English-proficient; SPED = special education; W = white. Measures: CAT = California Achievement Test; GMRT = Gates–MacGinitie Reading Tests; TOSREC = Test of Silent Reading Efficiency and Comprehension.

* $p < .05$.

(including teacher teams), and schoolwide approaches designed to make entire schools more focused on effective strategies for improving achievement and socioemotional development. We refer to these as whole-school models with organizational elements. Three studies of these two programs found a mean effect size of 0.09 ($p < .05$).

Five programs, each with one study, provided PD to all teachers of main academic subjects but did not emphasize organizational elements. These programs had a mean effect size of 0.00. The mean effect size across all eight studies of whole-school programs was 0.06 ($p < .05$). The evaluations of all seven programs are summarized in Table 5.

TABLE 5
Whole-School Approaches

Intervention	Design/ treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>Whole school with organizational elements</i>								
<i>Building Assets, Reducing Risks</i>								
T. Borman et al. (2017)	SR	1 year	Students: 2,172	9	71% minority; 21% EL; 70% FRL	NWEA	0.08*	0.08*
Corsello & Sharma (2015)	SR	1 year	Students: 495	9	52% W, 37% H, 11% AA; 17% EL; 68% FRL	NWEA	0.14*	0.14*
<i>Talent Development High School (Strategic Reading and Student Team Literature)</i>								
Balfanz et al. (2004)	CQE	1 year	Schools: 6 Teachers: 20 Students: 457	9	89% AA, 10% W; >90% FRL	TerraNova Comprehensive Test of Basic Skills	0.32	0.32
<i>Whole school without organizational elements</i>								
<i>Teacher Effectiveness Enhancement Programme</i>								
Institute for Effective Education (2016)	CR	1.5–2 years	Schools: 45 Students: 10,385	Year 9	United Kingdom; 16% EL; 30% FRL	Insight English (CEM)	-0.04	-0.04
<i>Chicago Striving Readers</i>								
Simon et al. (2011)	CR/TA	1–3 years	Schools: 59 Students: 8,127	6–8	58% AA, 35% H; 9% SPED; 96% FRL	Illinois Student Achievement Test Reading	-0.01	-0.01
<i>Project CRISS</i>								
Kushman et al. (2011)	CR	1 year	Schools: 49 Students: 4,959	9	79% W; 15% FRL	Stanford Diagnostic Reading Test	0.05	0.05
<i>eMINTS</i>								
Meyers et al. (2016)	CR/TA	3 years	Schools: 59 Students: 3,295	6–8	93% W	Measures of Academic Progress	-0.06	-0.06
<i>Texas Technology Immersion Pilot</i>								
Shapley et al. (2009)	CQE/TA	2–3 years	Schools: 42 Students: 10,234	6–8	70% H, 22% W, 7% AA; 15% LEP; 70% FRL	Texas Assessment of Knowledge and Skills	0.06	0.06

Note. Design/treatment: CQE = cluster quasi-experimental design; CR = cluster randomized; SR = student randomized; TA = technology application. Demographics: AA = African American; EL = English learner; FRL = receiving free or reduced-price lunch; H = Hispanic; LEP = limited English-proficient; SPED = special education; W = white. Measures: CEM = Centre for Evaluation and Monitoring; NWEA = Northwest Evaluation Association.

* $p < .05$.

Whole-School Approaches With Organizational Elements

There are two examples of whole-school approaches with organizational elements. One is Building Assets, Reducing Risks (BARR), a whole-school reform approach focused on developmental, academic, and structural challenges during ninth grade. BARR is used in all major subjects to attempt to increase students' achievement by improving their socioemotional skills, building positive student-teacher relationships, and solving nonacademic barriers to learning, such as truancy and behavioral problems. The strategy focuses on building students' personal assets and reducing substance abuse, delinquency, and other problems. BARR staff closely monitor student achievement, including real-time analysis of student data. Students take English, math, and science or social studies from an interdisciplinary team of teachers to build connections among students and teachers. Teachers in each team meet regularly to review the progress of at-risk students. Extensive PD and coaching are provided to teachers and school leaders. Corsello and Sharma (2015) found a positive effect of BARR on reading ($ES = 0.14, p < .01$), and a larger study by T. Borman et al. (2017) also found a statistically significant positive effect ($ES = 0.08, p < .05$). The weighted mean across the two studies was 0.09.

Another whole-school approach with organizational elements is Talent Development High School, a whole-school reform model for ninth graders. Within this program, Strategic Reading and Student Team Literature is the reading component. Students receive a double dose of reading and math, amounting to 90 minutes a day for each subject. In Strategic Reading and Student Team Literature, students work in small, interdependent cooperative learning groups using structured partner discussion guides that provide background and vocabulary, and comprehension questions provide minilessons on specific comprehension strategies.

Balfanz, Legters, and Jordan (2004) carried out an evaluation of Talent Development High School in high-poverty high schools in Baltimore, Maryland, comparing with control schools that also provided double-dose reading. There was a statistically nonsignificant effect size of 0.32 (significance was not attained because there were only six schools).

Whole-School Approaches Without Organizational Elements

Teacher Effectiveness Enhancement Programme is a U.K. example of a whole-school approach that provides PD to teachers of all major subjects but does not provide organizational elements beyond the classroom. The program focuses on formative assessment, thinking skills, cooperative learning, and effective use of technology.

An evaluation by the Institute for Effective Education (2016) found no positive effects ($ES = -0.04, n.s.$).

There was a clear distinction among whole-school approaches according to whether they included such organizational elements as teacher teams, schoolwide instructional emphases (e.g., cooperative learning), and socioemotional development. The two programs that did these things obtained positive outcomes. Those that provided PD to all teachers but did not focus on organizational change reported near-zero effect sizes. One explanation for the positive outcomes of BARR and Talent Development High School is that they focused on teacher-student and student-student relationships, not just on improving teachers' instructional skills. For adolescent learners, especially ones who are not succeeding in school, these relationships, and social motivations to learn and succeed, may be of particular importance.

Writing-Focused Approaches

Two programs, summarized in Table 6, focus on teaching writing as a means of improving reading outcomes. The weighted mean across the three studies of these programs was 0.13 ($p < .001$).

As one example, Pathway is a PD program used primarily with mainstreamed Latino English learners (ELs). Teachers learn how to teach cognitive strategies and process writing. Because of the focus on ELs, Pathway has a strong emphasis on vocabulary, relationships, and collaboration. A study by Olson et al. (2012) found statistically significant differences ($ES = 0.07, p < .05$), and the weighted average across the two Pathway studies (Olson et al., 2012; Olson, Matuchniak, Chung, Stumpf, & Farkas, 2016) was 0.08. Expository Reading and Writing Course is another example. It is a program for 12th graders designed to prepare them to pass the California Early Placement Test, used in the California State University system to determine whether freshmen must take noncredit remedial English courses. The emphasis of the program is on discussion of text meaning, developing critical thinking skills, encouraging group discussions, and developing writing skills in multiple genres. Fong, Finkelstein, Jaeger, Diaz, & Broek (2015) found statistically significant positive effects of the Expository Reading and Writing Course ($ES = 0.13, p < .05$).

Content-Focused Approaches

Seven studies evaluated content-focused approaches that teach comprehension strategies for social studies, science, and other expository text (see Table 7). Across all seven studies, the weighted mean effect size was 0.08 ($p < .05$).

An example of content-focused approaches is Reading Apprenticeship, a family of programs designed

TABLE 6
Writing-Focused Approaches

Intervention	Design/ treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>Pathway</i>								
Olson et al. (2012)	CR	1 year	Teachers: 161 Students: 4,459	6–11	95% H; 88% EL; 79% FRL	California Standards Test–English Language Arts	0.07*	0.07*
Olson et al. (2016)	CR	1 year	Teachers: 16 Students: 575	10	68% H, 18% A, 12% W; 20% EL; 71% FRL	California High School Exit Examination	0.19	0.19
<i>Expository Reading and Writing Course</i>								
Fong et al. (2015)	QE	1 year	Students: 6,618	12	45% H, 27% A, 24% W	Early Placement Test (California)	0.13*	0.13*

Note. Design/treatment: CR = cluster randomized; QE = Quasi-experimental design. Demographics: A = Asian; EL = English learner; FRL = receiving free or reduced-price lunch; H = Hispanic; W = white.

* $p < .05$.

to improve reading comprehension by integrating metacognitive strategy instruction into content areas, such as science and social studies. The model incorporates extensive reading, teaching of comprehension skills, writing, and collaborative sensemaking. Across four studies of Reading Apprenticeship variations, the average effect size was 0.07. As another example, Intelligent Tutoring System for the Structure Strategy is a web-based approach in which students are taught to comprehend nonfiction text by categorizing text structures using key elements in the text to find the main idea, activating prior knowledge, supporting cognitive monitoring, and using graphic organizers and flowcharts to summarize texts. Intelligent Tutoring System for the Structure Strategy is used during regular language arts classes with the assistance of a paraprofessional. Animated tutors model and guide learners. Students practice, take regular assessments, and proceed at their own pace through self-instructional units. A study by Wijekumar, Meyer, and Lei (2017) found a statistically significant positive effect for seventh graders ($ES = 0.18, p < .05$).

Vocabulary-Focused Approaches

Three programs emphasized the teaching of vocabulary as a means of enhancing comprehension. Findings of four studies of these programs, shown in Table 8, found a mean effect size of 0.06 (n.s.), corresponding to the outcomes on general comprehension reported by Wright and Cervetti (2017). For example, Academic Language Instruction for All Students is a vocabulary intervention designed to be used in regular English language

arts classrooms with many language-minority students. Each cycle of lessons is based on one informational text, from which are extracted a small number of high-utility and abstract words. The intervention includes a variety of whole-group, small-group, and independent activities. The weighted mean effect size across two studies of this program was 0.07 (n.s.; Lesaux, Kieffer, Faller, & Kelley, 2010; Lesaux, Kieffer, Kelley, & Harris, 2014).

Strategy-Focused Instruction

Strategy-focused instruction approaches are all variations of the SIM (Deshler & Schumaker, 2005), a family of programs that emphasize teaching students reading comprehension, decoding, and other reading objectives with step-by-step strategies. SIM studies, summarized in Table 9, had a mean effect size of 0.09 ($p < .001$).

SIM: Xtreme Reading is the main version of the SIM designed for struggling readers. Students reading two to five years below grade level are usually given an additional daily reading period, but in one of four studies (Faddis et al., 2011), Xtreme Reading was taught during an extra reading period in high schools but not in middle schools. Across the four qualifying studies (Faddis et al., 2011; Somers et al., 2010; Sprague, Zaller, Kite, & Hussar, 2012), the mean effect size for Xtreme Reading was 0.09 ($p < .01$).

Personalization Approaches

Personalization approaches provide content appropriate to students' reading levels, usually using computers. In some programs, there is a strong element of choice,

TABLE 7
Content-Focused Approaches

Intervention	Design/ treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>Reading Apprenticeship</i>								
Greenleaf et al. (2011)	CR	3 years	Schools: 78 Teachers: 111	9–11	48% H, 31% W; 19% EL; 41% FRL	Degree of Reading Power	–0.04	0.03
						California Standards Test–ELA	0.10	
						Reading comprehension	0.13	
<i>Reading Apprenticeship Improving Secondary Education</i>								
Fancsali et al. (2015)	CQE	1–2 years	Schools: 42 Students: 10,173	9–12	49% AA, 33% H; 10% EL; 40% FRL	ETS Literacy Assessment	0.14	0.14
<i>Internet-Based Reading Apprenticeship Improving Science Education</i>								
Jaciw et al. (2016)	CR	1 year	Schools: 26 Teachers: 69 Students: 1,468	9–12	73% W, 16% AA; 52% FRL	ETS Literacy Assessment	0.00	0.00
<i>Reading Apprenticeship Academic Literacy</i>								
Somers et al. (2010)	SR/ARP	1 year	Students: 2,255	9	31% H, 47% AA; 67% FRL	Group Reading Assessment and Diagnostic Evaluation	0.08	0.10*
						State tests ELA	0.15*	
<i>Every Classroom, Every Day/Literacy Matters</i>								
Early et al. (2016)	CR/ARP	2 years	Schools : 20 Students: 8,250	9 and 10	51% H, 24% AA; 22% EL; 76% FRL	State test ELA	0.06	0.06
<i>Intelligent Tutoring System for the Structure Strategy</i>								
Wijekumar et al. (2017)	CR/TA	6–7 months	Classrooms: 108 Students: 2,489	7	92% W, 8% minority; 42% FRL	Gray Silent Reading Test	0.18*	0.18*
<i>Content Knowledge-Building and Student-Regulated Comprehension Practices</i>								
Simmons et al. (2014)	CR	1 semester	Classes: 65 Students: 786	7–10	36% H, 31% AA; 71% FRL	Gates– MacGinitie Reading Tests	–0.01	–0.01

Note. Design/treatment: ARP = additional reading period; CQE = cluster quasi-experimental design; CR = cluster randomized; SR = student randomized; TA = technology application. Demographics: AA = African American; EL = English learner; FRL = receiving free or reduced-price lunch; H = Hispanic; W = white. Measures: ELA = English language arts; ETS: Educational Testing Service.

* $p < .05$.

where students may select texts of interest to them. Six studies, summarized in Table 10, fell into this category. The weighted mean effect size was 0.13 ($p < .05$).

As an example of personalization approaches, Success-Maker is an adaptive K–8 computer-based reading program. It provides individualized reading activities, game-like environments, interactive aids, and a reporting system to

inform teachers of student progress. Gatti (2011) found a statistically nonsignificant effect size for seventh graders of 0.11 (n.s.). Also, Accelerated Reader provides students with a wide range of books at their personal reading level, as determined by an online test. Online comprehension tests are provided for each book, and students can earn points based on completing many books at a high readability level. In a

TABLE 8
Vocabulary-Focused Approaches

Intervention	Design/ treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>Academic Language Instruction for All Students</i>								
Lesaux et al. (2010)	CR	18 weeks	Classes: 21 Students: 476	6	49% H; 73% EL	GMRT Comprehension	0.15	0.15
Lesaux et al. (2014)	CR	20 weeks	Teachers: 50 Students: 2,082	6	71% EL	GMRT	0.04	0.04
<i>Word Generation</i>								
Lawrence et al. (2017)	CR	1 year	Schools: 44 Students: 8,466	6–8	81% FRL	GMRT	0.05	0.05
<i>Vocabulary Enrichment Intervention Programme</i>								
Styles et al. (2014)	SR	6 months	Schools: 11 Students: 570	7	United Kingdom; 28% FRL	New Group Reading Test (United Kingdom)	0.06	0.06

Note. Design/treatment: CR = cluster randomized; SR = student randomized. Demographics: EL = English learner; FRL = receiving free or reduced-price lunch; H = Hispanic. Measures: GMRT = Gates–MacGinitie Reading Tests.

U.K. study, Gorard, Siddiqui, and See (2015) found a statistically significant effect size of 0.24 ($p < .05$).

Group/Personalization Rotation Approaches

Seventeen studies of nine programs, mostly using technology, rotated students through group instruction from the teacher, personalized work at students' own reading levels, and (often) cooperative learning. These studies, summarized in Table 11, had a weighted mean effect size of 0.09 ($p < .001$).

An example of group/personalization approaches is READ 180, an instructional model used 90 minutes each day with struggling readers. It entails 30 minutes of whole-group instruction, followed by one hour during which students rotate through three 20-minute blocks devoted to independent reading, small-group direct instruction with the teacher, and use of READ 180 adaptive software. READ 180 is always used in addition to ordinary English language arts, but in two studies, the control group also received supplemental instruction, so there was no difference in time. Across all five qualifying studies of READ 180 (Lang et al., 2009; Meisch et al., 2011; Schenck, Feighan, Coffey, & Rui, 2011; Sprague et al., 2012; Swanlund et al., 2012), the mean effect size was 0.09.

Intensive Group Approaches

Intensive group approaches are programs other than tutoring designed to catch students up on decoding

and word study skills, as well as basic comprehension strategies, which other students would have mastered in elementary school. These studies, summarized in Table 12, had a weighted mean effect size of 0.00 (n.s.). An example is REACH, which provides explicit, intensive instruction to struggling students in grades 6–12. It focuses on phonics, fluency, word knowledge, and reasoning skills; narrative structure and writing skills; and spelling. Lang et al. (2009) evaluated REACH and found a mean effect size of -0.02 (n.s.).

Cross-Cutting Factors

We included all studies in random effects models to explore two cross-cutting features that differed within treatment categories: extra reading periods and use of technology.

Extra Reading Periods

In many of the qualifying studies, the intervention was provided to groups of low achievers during an extra daily class reading period, always by certified teachers. The control group was typically participating in electives, such as art or band, or was in study hall, so the intervention provided substantial additional teaching time in reading to the experimental group over one or more years. In extra-time treatments, group sizes were typically small (usually 12–20 students).

TABLE 9
Strategy-Focused Instruction

Intervention	Design/treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>SIM: Xtreme Reading</i>								
Faddis et al. (2011) middle school	SR	1 year	Students: 822	7 and 8	34% H, 23% AA; 27% SPED; 34% EL	GRADE OAKS	0.29* 0.12	0.20*
Faddis et al. (2011) high school	SR/ARP	1 year	Students: 757	9 and 10	35% AA, 26% H; 24% SPED; 20% EL	GRADE OAKS (10th grade only)	0.12* 0.02	0.08
Somers et al. (2010)	SR/ARP	1 year	Students: 2,329	9	47% AA, 31% H; 67% FRL	GRADE State tests English language arts	0.04 0.08	0.06
Sprague et al. (2012)	SR/ARP	1 year	Students: 448	9	22% SPED; 75% FRL	Stanford Diagnostic Reading Test-4	0.00	0.00
<i>SIM: Content Literacy Curriculum</i>								
Corrin et al. (2012)	CR	1–2 years	Schools: 28 Students: 4,546	9 and 10	47% AA; 9% EL; 11% SPED; 65% FRL	GRADE	0.09	0.09
<i>SIM: Learning Strategies Curriculum</i>								
Cantrell et al. (2016)	SR/ARP	1 year	Grade 6 students: 1,135 Grade 9 students: 1,128	6 and 9	88% W; 26% SPED; 62% FRL	GRADE grade 6 GRADE grade 9	0.08 0.12*	0.10*
<i>SIM: Fusion Reading</i>								
Schiller et al. (2012)	SR/ARP	1 year	Students: 581	6–10	81% AA; 13% SPED	GRADE Michigan Educational Assessment Program Reading	0.05 0.11	0.07
<i>SIM: Adolescent Literacy Model</i>								
Cantrell et al. (2011)	CQE	4 years	Schools: 35	6–12	>90% W; 46% FRL	KCCT Reading middle schools KCCT Reading high schools	0.08 0.12	0.10

Note. Design/treatment: ARP = additional reading period; CQE = cluster quasi-experimental design; CR = cluster randomized; SR = student randomized. Demographics: AA = African American; EL = English learner; FRL = receiving free or reduced-price lunch; H = Hispanic; SPED = special education; W = white. Measures: GRADE = Group Reading Assessment and Diagnostic Evaluation; KCCT = Kentucky Core Content Test; OAKS = Oregon Assessment of Knowledge and Skills.

* $p < .05$.

In an analysis comparing studies providing extra reading periods and studies that did not, we found no statistically significant differences. Twenty-seven studies of programs

providing extra reading time had a mean weighted effect size of 0.09, whereas 42 studies of programs providing no extra time had a weighted mean effect size of 0.10.

TABLE 10
Personalization Approaches

Intervention	Design/ treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>Thinking Reader</i>								
Drummond et al. (2011)	CR	1 year	Teachers: 90 Students: 2,149	6	37% W, 28% H; 11% SPED; 10% EL; 71% FRL	GMRT	0.01	0.01
<i>Schoolwide Enrichment Model-Reading</i>								
Little et al. (2014)	CR	1 year	Teachers: 47 Students: 2,028	6-8	48% AA, 32% H; 67% FRL	GMRT	0.10	0.10
<i>Achieve3000</i>								
Shannon & Grant (2015)	CR	1 year	Teachers: 33 Grade 6 students: 494 Grade 9 students: 248	6 and 9	67% W, 37% H; 12% SPED; 12% EL; 62% FRL	GMRT	0.29*	0.29*
						Grade 6	0.22	
						Grade 9	0.44	
<i>SuccessMaker</i>								
Gatti (2011)	CR	1 year	Classes: 22 Students: 453	7	51% W; 53% FRL	GRADE	0.11	0.11
<i>Accelerated Reader</i>								
Gorard et al. (2015)	SR/TA	22 weeks	Students: 349	Year 7	United Kingdom; 88% W; 23% SPED; 35% FRL	New Group Reading Test (United Kingdom)	0.24*	0.24*
<i>iLit</i>								
Gatti (2016)	SR/ARP	2 years	Students: 213	7 and 8	53% H, 22% W, 17% AA; 26% LEP; 13% SPED; 80% FRL	GRADE	0.09	0.09

Note. Design/treatment: ARP = additional reading period; CR = cluster randomized; SR = student randomized; TA = technology application. Demographics: AA = African American; EL = English learner; FRL = receiving free or reduced-price lunch; H = Hispanic; LEP = limited English-proficient; SPED = special education; W = white. Measures: GMRT = Gates-MacGinitie Reading Tests; GRADE = Group Reading Assessment and Diagnostic Evaluation.

* $p < .05$.

Use of Technology

We compared programs making extensive use of technology with those that made little or no use of technology. Programs using technology ($n = 23$) obtained statistically nonsignificantly lower effect sizes ($ES = 0.08$) than those that did not use technology ($n = 46$; $ES = 0.10$). A few individual programs making extensive use of technology, such as Achieve3000 ($ES = 0.29$), Accelerated Reader ($ES = 0.24$), and Intelligent Tutoring System for the Structure Strategy ($ES = 0.18$) reported statistically significant positive impacts, but these were exceptions.

Demographic and Methodological Moderators

Struggling Readers and ELs

Some qualifying studies served only struggling readers, usually those performing at least two years below grade level. Other studies served all students, although in most cases, the schools involved were high in poverty. We compared outcomes for studies focused only on struggling readers ($n = 36$) with those focused on all students ($n = 36$). Effect sizes were nearly identical: 0.10 for struggling readers and 0.10 for all students.

TABLE 11
Group/Personalization Rotation Approaches

Intervention	Design/ treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>READ 180</i>								
Schenck et al. (2011)	SR/ARP	2 years	Students: 1,295	6–8	93% AA; 6% EL; 92% FRL	Iowa Test of Basic Skills	0.02	0.02
Meisch et al. (2011)	CR	3 years	Schools: 19 Students: 1,023	6–8	55% AA, 42% H; 13% EL; 44% SPED; 62% FRL	Stanford Achievement Test 10	0.06	0.06
Swanlund et al. (2012)	SR/ARP	1 year	Students: 619	6–9	70% AA, 19% H; 36% SPED; 8% EL; 88% FRL	Measures of Academic Progress–Reading	0.14*	0.14*
Lang et al. (2009)	SR	1 year	High-risk students: 190 Moderate-risk students: 409	9	19% AA, 19% H; 41% FRL	FCAT Reading High risk Moderate risk	–0.27* 0.30*	0.12
Sprague et al. (2012)	SR/ARP	1 year	Students: 456	9	73% minority; 19% SPED; 72% FRL	Stanford Diagnostic Reading Test–4	0.18*	0.18*
<i>Expert 21 (related to READ 180)</i>								
Sivin-Kachala & Bialo (2012)	CR/TA	1 year	Teachers: 6 Students: 276	6–8	71% H, 27% AA; 100% FRL	New Jersey Assessment of Skills and Knowledge GMRT	0.20 0.10	0.15
<i>System 44 (related to READ 180)</i>								
Beam et al. (2011)	SR/ARP/TA	1 year	Students: 147	6–8	63% W; 7% SPED; 33% FRL	TOSREC California Standards Test	–0.24* –0.04	–0.14
Beam & Faddis (2012)	SR/ARP/TA	1 year	Students: 145	6–8	78% AA; 53% SPED; 96% FRL	TOSREC	0.20*	0.20*
<i>Passport Reading Journeys</i>								
Vaden-Kiernan et al. (2012)	SR/ARP/TA	1 year	Students: 1,102	6 and 7	76% minority; 15% SPED; 88% FRL	Integrated Louisiana Educational Assessment Program–Reading GRADE	–0.01 0.27*	0.12*
Schenck et al. (2012)	SR/ARP/TA	1 year	Students: 701	7 and 8	68% AA; 24% SPED; 8% EL; 88% FRL	Standards of Learning (English/reading, Virginia)–Reading GMRT Overall Comprehension Vocabulary	0.06 0.06 0.05 0.07	0.06
Dimitrov et al. (2012)	SR/ARP/TA	1 year	Students: 460	9	58% AA, 30% W, 5% H; 18% SPED; 85% FRL	GMRT Explore	0.02 –0.09	–0.03

(continued)

TABLE 11
Group/Personalization Rotation Approaches (continued)

Intervention	Design/ treatment	Duration	N	Grade(s)	Sample characteristics	Posttest	Effect size	Overall effect size
<i>Comprehension Circuit Training</i>								
Fogarty et al. (2014)	CR/TA	1 semester	Classes: 61 Students: 859	6–8	43% H, 35% W, 22% AA; 9% EL; 6% SPED; 67% FRL	GMRT	0.12	0.12
Fogarty et al. (2017)	CR/TA	1 semester	Classes: 16 Students: 228	6–8	30% AA, 27% W, 26% H; 9% SPED; 67% FRL	State of Texas Assessments of Academic Readiness	0.09	0.15
						GMRT	0.12	
						GRADE	0.11	
						Gray Oral Reading Test	0.18	
<i>Prentice Hall Literature</i>								
Eddy et al. (2010)	CR/TA	1 year	Teachers: 29 Students: 1,518	7–10	55% H, 15% AA	GMRT	–0.10	–0.10
<i>Strategic Adolescent Reading Intervention</i>								
Kim et al. (2017)	SR/ARP	1 year	Schools: 8 Students: 398	6–8	49% W, 26% H, 19% AA; 13% EL; 69% FRL	Reading Inventory and Scholastic Evaluation	0.15	0.15
<i>Read to Achieve</i>								
Deussen et al. (2012)	SR/ARP	1 year	Students: 325	6 and 8	43% W; 23% EL; 58% FRL	GMRT	0.02	0.10
						Measurements of Student Progress (state reading tests)	0.16	
<i>Reading Intervention through Strategy Enhancement</i>								
Lang et al. (2009)	SR	1 year	High-risk students: 194 Moderate- risk students: 400	9	19% AA, 19% H; 43% FRL	FCAT Reading		0.16*
						High risk	–0.06	
						Moderate risk	0.27*	

Note. Design/treatment: ARP = additional reading period; CR = cluster randomized; SR = student randomized; TA = technology application. Demographics: AA = African American; EL = English learner; FRL = receiving free or reduced-price lunch; H = Hispanic; SPED = special education; W = white. Measures: FCAT = Florida Comprehensive Assessment Test; GMRT = Gates–MacGinitie Reading Tests; GRADE = Group Reading Assessment and Diagnostic Evaluation; TOSREC = Test of Silent Reading Efficiency and Comprehension.

* $p < .05$.

None of the qualifying studies reported outcomes separately for ELs, but two programs focused on ELs and showed promising outcomes: Pathway (mean ES = 0.08) and Academic Language Instruction for All Students (mean ES = 0.06). Among the two studies of Pathway, one found statistically significant positive effects.

Middle Versus High School

We tested the difference in outcomes between programs used in the middle grades (6–8) and those used in high

school (9–12). It is important to note that most of the high school studies focused on grades 9–10. Weighted mean effect sizes were nearly identical in middle schools ($n = 44$ studies; ES = 0.10) and in high schools ($n = 30$ studies; ES = 0.10).

Differences by Research Design

We compared effect sizes between studies that used random assignment to conditions ($n = 62$ studies; ES = 0.09) and quasi-experiments, which used matching

TABLE 12
Intensive Group Approaches

Intervention	Design/ treatment	Duration	N	Grade	Sample characteristics	Posttest	Effect size	Overall effect size
<i>Reading Excellence: Word Attack and Rate Development Strategies</i>								
Newman et al. (2012)	SR/ARP	1 year	Students: 517	7	64% H, 22% AA, 12% A; 95% FRL	New York State English Language Arts Gates–MacGinitie Reading Tests Total	0.15 0.02	0.09
<i>Kentucky Cognitive Literacy Model</i>								
Cantrell et al. (2012)	SR/ARP	1 year	Students: 485	9	88% W; 16% SPED; 62% FRL	Group Reading Assessment and Diagnostic Evaluation	-0.06	-0.06
<i>REACH</i>								
Lang et al. (2009)	SR	1 year	High-risk students: 181 Moderate-risk students: 401	9	20% AA, 20% H; 43% FRL	Florida Comprehensive Assessment Test–Reading High risk Moderate risk	 -0.19 0.06	-0.02

Note. Design/treatment: ARP = additional reading period; SR = student randomized. Demographics: A = Asian; AA = African American; FRL = receiving free or reduced-price lunch; H = Hispanic; SPED = special education; W = white.

($n = 7$ studies, $ES = 0.13$). This difference was statistically nonsignificant but similar to differences reported by Cheung and Slavin (2016). Effect sizes were statistically nonsignificantly lower for studies using clustered designs ($n = 34$; $ES = 0.05$) than for those using designs employing student-level analyses ($n = 35$; $ES = 0.10$).

Programs Meeting ESSA Standards for Strong and Moderate Evidence of Effectiveness

The ESSA defines “strong” and “moderate” criteria for evidence supporting educational programs. “Strong” requires that at least one randomized study found statistically significant positive effects and no statistically significant negative effects, and “moderate” requires that at least one quasi-experimental (matched) study found statistically significant positive effects and no statistically significant negative effects.

Table 13 summarizes the programs that met these ESSA categories, along with the numbers of studies, weighted mean effect sizes, and ESSA ratings. What the table illustrates is that even though only a few categories are particularly associated with positive outcomes, most categories contain at least one individual program that has been found in at least one study to have statistically significant positive outcomes on secondary reading measures.

Discussion

In this review of rigorous research on programs designed to enhance the reading of students in middle and high schools, we found that studies meeting inclusion criteria had a wide range of effects on student reading. As noted earlier, the mean effect size of 0.19 among statistically significant studies could be used to reduce reading gaps between minority and white students by about 38%. Weighted mean effect sizes for all categories are shown in Table 2. The table also shows a lack of statistically significant differences for two cross-cutting factors (extra reading time and use of technology), as well as four demographic and methodological moderators (struggling readers/all students, middle/high school, randomized/quasi-experimental, and clustered/student-level).

A few quite different categories contained programs with more positive impacts. One of these was tutoring programs (all done in the United Kingdom), in which teachers or paraprofessionals worked with groups of one to four students (or in one case, two adults to six to eight students). The weighted mean effect size across all five studies was 0.24. It is not surprising that tutoring would be effective, as it has also been effective in elementary reading (Inns et al., 2018; Slavin et al., 2011; Wanzek et al., 2013). Tutors are able to thoroughly personalize instruction to students’ individual needs and to build personal relationships with

TABLE 13
Programs Meeting ESSA Evidence Standards for “Strong” and “Moderate” Ratings

Program	Number of studies	Average effect size	ESSA rating
<i>Tutoring</i>			
Perry Beeches Coaching Programme	1	0.36	Strong
REACH Tutoring	1	0.42	Strong
Butterfly Phonics	1	0.30	Strong
<i>Cooperative learning</i>			
The Reading Edge	3	0.29	Strong
<i>Whole-school approach</i>			
Building Assets, Reducing Risks	2	0.09	Strong
<i>Writing-focused approaches</i>			
Pathway	2	0.08	Strong
Expository Reading and Writing Course	1	0.13	Moderate
<i>Content-focused approaches</i>			
Reading Apprenticeship	4	0.10	Strong
Intelligent Tutoring System for the Structure Strategy	1	0.18	Strong
<i>Strategy instruction</i>			
SIM	7	0.09	Strong
<i>Personalization approaches</i>			
Achieve3000	1	0.29	Strong
Accelerated Reader	1	0.24	Strong
<i>Group/personalization rotation approaches</i>			
READ180	5	0.08	Strong
Passport Reading Journeys	3	0.07	Strong
Comprehension Circuit Training	2	0.13	Strong
Strategic Adolescent Reading Intervention	1	0.15	Strong
Reading Intervention through Strategy Enhancement	1	0.16	Strong

students. However, single studies of cross-age peer tutoring and volunteer tutoring approaches did not report positive outcomes.

Another category with particularly positive outcomes was cooperative learning, especially The Reading Edge, with one study with a statistically significant positive outcome and a weighted mean effect size across three studies of 0.15. What makes cooperative learning distinctive is that it taps into the social motivations that drive much of adolescent behavior. By having students work in teams, with team recognition based on the achievement gains of all team members, teammates encourage one another’s efforts, explain ideas to one

another, and have opportunities to ask others for help (Guthrie, 2015; Roseth et al., 2008; Slavin, 1995).

Our conclusion that tutoring and cooperative learning were effective approaches in secondary reading is consistent with the conclusions of previous reviews by Slavin et al. (2008), Dietrichson et al. (2017), and Herrera et al. (2016).

Three studies of whole-school designs with organizational elements such as teacher teams showed particular promise. BARR, a whole-school model, has been evaluated in two high-quality randomized experiments. BARR organizes ninth-grade teachers in interdisciplinary teams that share responsibility for all aspects of

student development in high-poverty high schools. Teachers meet individually with students to plan and review progress toward the students' own goals. The program emphasizes socioemotional development and relationships between teachers and students. A study of Talent Development High School, which also is a whole-school approach with teacher teaming, had an impressive but statistically nonsignificant effect size of 0.32. These approaches may show the benefits of focusing on the socioemotional development of young adolescents. In contrast, other whole-school approaches lacking organizational elements designed to enhance socioemotional development did not show positive outcomes.

Two programs with a strong emphasis on writing found positive effects on reading. The Expository Reading and Writing Course, a program designed to help 12th graders prepare for the test they will take as freshmen if they attend college in the California State University system, reported a statistically significant positive effect on this test, with an effect size of 0.13. Pathway, a PD program primarily for teaching mainstreamed ELs, also showed statistically significant effects across two studies ($ES = 0.08$). Writing-focused programs are likely to have improved reading outcomes because extensive writing is believed to help students build understanding of the author's perspective by becoming authors themselves (Graham et al., 2017; Graham & Hebert, 2011). It may also be important that these two particular programs both used writing for well-defined purposes, one (Expository Reading and Writing Course) preparing 12th graders for college and the other (Pathway) improving English among ELs. The mean for this category was 0.13.

One family of approaches with mostly positive effects was the SIM, which teaches struggling adolescent readers step-by-step strategies for phonics, comprehension, writing, note-taking, and other skills. Eight studies evaluated SIM variations, and three of these found statistically significant positive effects. These approaches depend on making young adolescents strategic, autonomous comprehenders, using proven strategies (Frankel et al., 2016). The outcomes of all eight studies averaged 0.09.

Another family of programs with a strong focus on teaching specific strategies is Reading Apprenticeship (Greenleaf et al., 2011). All four variations focus on comprehension and writing strategies. The only variation to have statistically significant positive effects was Reading Apprenticeship Academic Literacy (RAAL; Somers et al., 2010), designed for struggling readers. The mean across the four studies was 0.07.

One of the most surprising findings of our review is the lack of positive effects of providing an additional class period for reading each day. Programs focused on improving teachers' practices during regular class periods produced virtually identical impacts as programs that also focused on improving teachers' practices but

added a daily period of instruction. The finding matches the conclusions of reviews by Kidron and Lindsay (2014) and Zief et al. (2006).

The failure to find any impact of additional instructional time in reading was not expected. It seems obvious that an entire additional reading period each day would benefit student learning. Perhaps the problem is that struggling readers were unhappy about having to take a remedial reading class (instead of art, music, or physical education, in most cases) and were not motivated to once again work on material that they had difficulty with in elementary school.

Similarly, it was surprising to find that there was no impact of programs emphasizing technology, although there were a few individual exceptions. An earlier review by Cheung and Slavin (2013) found mostly positive, although small, outcomes of technology for reading outcomes, but a review of research on programs for struggling readers in elementary schools found no program to have a statistically significant positive effect (Inns et al., 2018). It may be that teachers are still not comfortable with technology. One indication of this might be that among the few technology-focused programs for secondary schools that had positive impacts, such as READ 180, Passport Reading Journeys, Achieve3000, and Accelerated Reader, all were older, well-established approaches that teachers may have found easier to implement.

There are a few commonalities among programs that achieved positive outcomes. Several of these are worth mentioning, although they are far from conclusive.

One interesting commonality was that programs with positive outcomes tended to emphasize student motivation, student-to-student and student-to-teacher relationships, and socioemotional learning (Guthrie, 2015). Positive examples include cooperative learning (at least The Reading Edge), BARR, and Talent Development High School, which focus on relationships and socioemotional learning, and tutoring, which provides immediate feedback and potentially close teacher-student relationships.

The positive impacts of both writing-focused approaches—Expository Reading and Writing Course and Pathway—replicate earlier research summarized by Graham and Hebert (2011). Learning to write well may help students gain insight into the structure of text, as they learn authors' tricks of the trade by being authors themselves.

It is clear that successful programs can appear in many categories. As research in secondary reading continues, it will be useful and instructive to vary program components to learn which generic approaches most enhance student outcomes. Yet, as Table 13 makes clear, programs currently meeting ESSA evidence standards are found in most categories, as are programs that do not meet ESSA standards.

There is still a need for much more research on secondary reading approaches. Many of the most promising categories were represented by just a few studies meeting review criteria: Tutoring by teachers and aides (5), cooperative learning (6), whole-school models using organizational elements (2), and writing-focused approaches (2) seem particularly likely to yield continued positive findings, and it is critical to find out which types of models focused on content, vocabulary, strategy teaching, and personalization have the greatest potential for secondary learning. Further research on approaches for struggling learners and ELs is of particular importance. Understandings of how to use powerful digital technologies is especially needed: Technology has a certain place in the future of education, but up to the present, its outcomes have been uneven.

The research reviewed here identified specific proven programs and outlined promising avenues toward more effective approaches, but much remains to be done to understand how to create replicable, cost-effective strategies that can reliably and meaningfully improve reading outcomes for middle and high school students.

NOTES

This article was written under partial funding from the Spencer Foundation (grant 201600014). However, any opinions expressed are those of the authors and do not represent the positions or policies of our funders. We disclose that Robert Slavin is a cofounder and former chairman of the nonprofit Success for All Foundation, which created the program The Reading Edge.

REFERENCES

References marked with an asterisk indicate studies included in the review.

- *Balfanz, R., Legters, N., & Jordan, W. (2004). *Catching up: Impact of the Talent Development ninth grade instructional interventions in reading and mathematics in high-poverty high schools* (Report No. 69). Baltimore, MD: Center for Research on the Education of Students Placed at Risk, Johns Hopkins University
- Beach, R., & O'Brien, D. (2018). Significant literacy research informing English language arts instruction. In D. Lapp & D. Fisher (Eds.), *Handbook of research on teaching the English language arts* (4th ed., pp. 1–30). New York, NY: Routledge.
- *Beam, M., & Faddis, B. (2012). *Evaluation of System 44: Final report*. Portland, OR: RMC Research.
- *Beam, M., Faddis, B., & Hahn, K. (2011). *Evaluation of System 44: Final report*. Portland, OR: RMC Research
- Biancarosa, G., & Snow, C.E. (2006). *Reading next—A vision for action and research in middle and high school literacy* (2nd ed.). Washington, DC: Alliance for Excellent Education.
- Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.
- Borman, G.D., Hewes, G.M., Overman, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73(2), 125–230. <https://doi.org/10.3102/00346543073002125>
- Borman, T., Bos, J., O'Brien, B., Park, S., Liu, F., Corsello, M., & Jerabek, A. (2017). *The BARR program: Impacting social emotional skills and academic achievement of 9th grade students in 6 high schools*. Washington, DC: American Institutes for Research.

- Boulay, B., Goodson, B., Frye, M., Blocklin, M., & Price, C. (2015). *Summary of research generated by Striving Readers on the effectiveness of interventions for struggling adolescent readers* (NCEE-2016-4001). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- *Cantrell, S.C., Almasi, J.F., Carter, J.C., & Rintamaa, M. (2011). *Striving Readers final evaluation report: Danville, Kentucky*. Lexington: Collaborative Center for Literacy Development, University of Kentucky.
- *Cantrell, S.C., Almasi, J.F., Rintamaa, M., & Carter, J.C. (2016). Supplemental reading strategy instruction for adolescents: A randomized trial and follow-up study. *The Journal of Educational Research*, 109(1), 7–26. <https://doi.org/10.1080/00220671.2014.917258>
- *Cantrell, S.C., Carter, J.C., & Rintamaa, M. (2012). *Striving Readers cohort II evaluation report: Kentucky*. Lexington: Collaborative Center for Literacy Development, University of Kentucky.
- Cheung, A., & Slavin, R.E. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly*, 48(3), 277–299. <https://doi.org/10.1002/rrq.50>
- Cheung, A., & Slavin, R.E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, 45(5), 283–292. <https://doi.org/10.3102/0013189X16656615>
- Cooper, H. (1998). *Synthesizing research* (3rd ed.). Thousand Oaks, CA: Sage.
- *Corrin, W., Lindsay, J.J., Somers, M.-A., Myers, N.E., Meyers, C.V., Condon, C.A., & Smith, J.K. (2012). *Evaluation of the Content Literacy Continuum: Report on program impacts, program fidelity, and contrast* (NCEE 2013-4001). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Corsello, M., & Sharma, A. (2015). *The Building Assets-Reducing Risks Program: Replication and expansion of an effective strategy to turn around low-achieving schools: i3 Development Grant: Final Report*. Retrieved from ERIC database. (ED560804)
- de Boer, H., Donker, A.S., & van der Werf, M.P.C. (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research*, 84(4), 509–545. <https://doi.org/10.3102/0034654314540006>
- Denver Public Schools. (2016). *Final report to the U.S. Department of Education on Collaborative Strategic Reading*. Denver, CO: Author.
- Deshler, D.D., & Schumaker, J.B. (Eds.) (2005). *Teaching adolescents with disabilities: Accessing the general education curriculum*. Thousand Oaks, CA: Corwin.
- Desimone, L.M., & Garet, M.S. (2015). Best practices in teacher's professional development in the United States. *Psychology, Society, & Education*, 7(3), 252–263.
- *Deussen, T., Scott, C., Nelsestuen, K., Roccograndi, A., & Davis, A. (2012). *Washington Striving Readers: Year 1 evaluation report*. Portland, OR: Education Northwest
- Dietrichson, J., Bøg, M., Filges, T., & Jørgensen, A.-M.K. (2017). Academic interventions for elementary and middle school students with low socioeconomic status: A systematic review and meta-analysis. *Review of Educational Research*, 87(2), 243–282. <https://doi.org/10.3102/0034654316687036>
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231–264. <https://doi.org/10.1007/s11409-008-9029-x>
- *Dimitrov, D., Jurich, S., Frye, M., Lammert, J., Sayko, S., & Taylor, L. (2012). *Year one evaluation report/impact study: Illinois Striving Readers*. Arlington, VA: RMC Research.
- Dougherty, S.M. (2013, March). *Adolescent literacy on the margin: Regression-discontinuity evidence from a “double dose” middle*

- grades literacy intervention. Paper presented at the annual meeting of the Society for Research on Educational Effectiveness, Washington, DC.
- *Drummond, K., Chinen, M., & Duncan, T.G., Miller, H.R., Fryer, L., & Zmach, C. (2011). *Impact of the Thinking Reader® software program on grade 6 reading vocabulary, comprehension, strategies, and motivation: Final report* (NCEE 2010-4035). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- *Early, D., Berg, J.K., Alicea, S., Si, Y., Aber, J.L., Ryan, R.M., & Deci, E.L. (2016). The impact of Every Classroom, Every Day on high school student achievement: Results from a school-randomized trial. *Journal of Research on Educational Effectiveness*, 9(1), 3–29. <https://doi.org/10.1080/19345747.2015.1055638>
- *Eddy, R.M., Ruitman, H.T., Hankel, N., & Sloper, M. (2010). *The effects of Pearson Prentice Hall Literature on student performance: Efficacy study final report*. La Verne, CA: Cobblestone.
- Edmonds, M.S., Vaughn, S., Wexler, J., Reutebuch, C.K., Cable, A., Tackett, K.K., & Schnakenberg, J.W. (2009). A synthesis of reading interventions and effects on reading outcomes for older struggling readers. *Review of Educational Research*, 79(1), 262–300. <https://doi.org/10.3102/0034654308325998>
- Every Student Succeeds Act of 2015, Pub. L. No. 114-95, § 114, Stat. 1177 (2016).
- *Faddis, B., Beam, M., Maxim, L., Vale Gandhi, E., Hahn, K., & Hale, R. (2011). *Portland Public Schools' Striving Readers program: Year 5 evaluation report*. Portland, OR: RMC Research.
- *Fancsali, C., Abe, Y., Pyatigorsky, M., Ortiz, L., Chan, V., Saltares, E., ... Jaciw, A. (2015). *The impact of the Reading Apprenticeship Improving Secondary Education (RAISE) project on academic literacy in high school: A report of a randomized experiment in Pennsylvania and California schools*. Palo Alto, CA: Empirical Education.
- Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, 35(1), 39–50. https://doi.org/10.1207/S15326985EP3501_5
- Flynn, L.J., Zheng, X., & Swanson, H.L. (2012). Instructing struggling older readers: A selective meta-analysis of intervention research. *Learning Disabilities Research & Practice*, 27(1), 21–32. <https://doi.org/10.1111/j.1540-5826.2011.00347.x>
- *Fogarty, M., Clemens, N., Simmons, D., Anderson, L., Davis, J., Smith, A., Oslund, E. (2017). Impact of a technology-mediated reading intervention on adolescents' reading comprehension. *Journal of Research on Educational Effectiveness*, 10(2), 326–353. <https://doi.org/10.1080/19345747.2016.1227412>
- *Fogarty, M., Oslund, E., Simmons, D., Davis, J., Simmons, L., Anderson, L., Roberts, G. (2014). Examining the effectiveness of a multicomponent reading comprehension intervention in middle schools: A focus on treatment fidelity. *Educational Psychology Review*, 26(3), 425–449. <https://doi.org/10.1007/s10648-014-9270-6>
- Fong, A.B., Finkelstein, N.D., Jaeger, L.M., Diaz, R., & Broek, M.E. (2015). *Evaluation of the Expository Reading and Writing Course: Findings from the Investing in Innovation development grant*. San Francisco, CA: WestEd.
- Frankel, K.K., Becker, B.L.C., Rowe, M.W., & Pearson, P.D. (2016). From “what is reading?” to what is literacy? *Journal of Education*, 196(3), 7–17. <https://doi.org/10.1177/002205741619600303>
- Frankel, K.K., Jaeger, E.L., Brooks, M.D., & Randel, M.A. (2015). Struggling readers: Using theory to complicate misunderstandings of what it means to be literate in school. *Literacy Research: Theory, Method, and Practice*, 64(1), 307–322. <https://doi.org/10.1177/2381336915617606>
- *Gatti, G.G. (2011). *Pearson SuccessMaker reading efficacy study: 2010–11 final report*. Pittsburgh, PA: Gatti Evaluation.
- *Gatti, G.G. (2016). *iLit 2013–15 longitudinal efficacy study: Final report*. Pittsburgh, PA: Gatti Evaluation
- Goldman, S.R. (2012). Adolescent literacy: Learning and understanding content. *The Future of Children*, 22(2), 89–116. <https://doi.org/10.1353/foc.2012.0011>
- *Gorard, S., Siddiqui, N., & See, B.H. (2015). *Accelerated Reader: Evaluation report and executive summary*. London, UK: Education Endowment Foundation.
- Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81(4), 710–744. <https://doi.org/10.17763/haer.81.4.t2k0m13756113566>
- Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K.R., & Holzapfel, J. (2017). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly*, 53(3), 279–304. <https://doi.org/10.1002/rrq.194>
- *Greenleaf, C.L., Litman, C., Hanson, T.L., Rosen, R., Boscardin, C.K., Herman, J., ... Jones, B. (2011). Integrating literacy and science in biology: Teaching and learning impacts of Reading Apprenticeship professional development. *American Educational Research Journal*, 48(3), 647–717. <https://doi.org/10.3102/0002831210384839>
- Guthrie, J.T. (2015). Growth of motivations for cognitive processes of reading. In P.D. Pearson & E.H. Hiebert (Eds.), *Research-based practices for teaching Common Core literacy* (pp. 107–122). New York, NY: Teachers College Press.
- Guthrie, J.T., Klauda, S.L., & Ho, A.N. (2013). Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. *Reading Research Quarterly*, 48(1), 9–26. <https://doi.org/10.1002/rrq.035>
- Hedges, L.V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, 32(4), 341–370. <https://doi.org/10.3102/1076998606298043>
- Herrera, S., Truckenmiller, A.J., & Foorman, B.R. (2016). *Summary of 20 years of research on the effectiveness of adolescent literacy programs and practices* (REL 2016-178). Washington, DC: Regional Educational Laboratory Southeast, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Hodges, T.S., Feng, L., Kuo, L.-J., & McTigue, E. (2016). Discovering the literacy gap: A systematic review of reading and writing theories in research. *Cogent Education*, 3, article 1228284. <https://doi.org/10.1080/2331186x.2016.1228284>
- Inns, A., Lake, C., Pellegrini, M., & Slavin, R.E. (2018, March). *Effective programs for struggling readers: A best-evidence synthesis*. Paper presented at the annual meeting of the Society for Research on Educational Effectiveness, Washington, DC.
- *Institute for Effective Education (2016). *Teacher Effectiveness Enhancement Programme: Evaluation report and executive summary*. York, UK: Author.
- *Jaciw, A.P., Schellinger, A.M., Lin, L., Zacamy, J., & Toby, M. (2016). *Effectiveness of Internet-based Reading Apprenticeship Improving Science Education (iRAISE): A report of a randomized experiment in Michigan and Pennsylvania*. Palo Alto, CA: Empirical Education.
- Jun, S.W., Ramirez, G., & Cumming, A. (2010). Tutoring adolescents in literacy: A meta-analysis. *McGill Journal of Education*, 45(2), 219–238. <https://doi.org/10.7202/045605ar>
- Kamil, M.L., Borman, G.D., Dole, J., Kral, C.C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE 2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Kervin, L., Mantei, J., & Leu, D.J. (2018). Repositioning online reading to a central location in the language arts. In D. Lapp & D.

- Fisher (Eds.), *Handbook of research on teaching the English language arts* (4th ed., pp. 327–359). New York, NY: Routledge.
- Kidron, Y., & Lindsay, J. (2014). *The effects of increased learning time on student academic and nonacademic outcomes: Findings from a meta-analytic review* (REL 2014-015). Washington, DC: Regional Educational Laboratory Appalachia, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- *Kim, J.S., Hemphill, L., Troyer, M., Thomson, J.M., Jones, S.M., LaRusso, M.D., & Donovan, S. (2017). Engaging struggling adolescent readers to improve reading skills. *Reading Research Quarterly*, 52(3), 357–382. <https://doi.org/10.1002/rrq.171>
- *King, B., & Kasim, A. (2015). *Rapid Phonics: Evaluation report and executive summary*. London, UK: Educational Endowment Foundation.
- *Kushman, J., Hanita, M., & Raphael, J. (2011). *An experimental study of the Project CRISS reading program on grade 9 reading achievement in rural high schools: Final report* (NCEE 2011-4007). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education
- *Lang, L., Torgesen, J., Vogel, W., Chanter, C., Lefsky, E., & Petscher, Y. (2009). Exploring the relative effectiveness of reading interventions for high school students. *Journal of Research on Educational Effectiveness*, 2(2), 149–175. <https://doi.org/10.1080/19345740802641535>
- Langer, J.A. (2001). Beating the odds: Teaching middle and high school students to read and write well. *American Educational Research Journal*, 38(4), 837–880. <https://doi.org/10.3102/00028312038004837>
- *Lawrence, J.F., Francis, D., Paré-Blagoev, J., & Snow, C.E. (2017). The poor get richer: Heterogeneity in the efficacy of a school-level intervention for academic language. *Journal of Research on Educational Effectiveness*, 10(4), 767–793. <https://doi.org/10.1080/19345747.2016.1237596>
- Lee, C.D., & Sprattley, A. (2010). *Reading in the disciplines: The challenges of adolescent literacy*. New York, NY: Carnegie Corporation of New York.
- *Lesaux, N.K., Kieffer, M.J., Faller, S.E., & Kelley, J.G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*, 45(2), 196–228. <https://doi.org/10.1598/RRQ.45.2.3>
- *Lesaux, N.K., Kieffer, M.J., Kelley, J.G., & Harris, J.R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, 51(6), 1159–1194. <https://doi.org/10.3102/0002831214532165>
- Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- *Little, C.A., McCoach, D.B., & Reis, S.M. (2014). Effects of differentiated reading instruction on student achievement in middle school. *Journal of Advanced Academics*, 25(4), 384–402. <https://doi.org/10.1177/1932202X14549250>
- *Lloyd, C., Edovald, T., Kiss, Z., Morris, S., Skipp, A., & Ahmed, H. (2015). *Paired Reading: Evaluation report and executive summary*. London, UK: Education Endowment Foundation.
- *Lord, P., Bradshaw, S., Stevens, E., & Styles, B. (2015). *Perry Beeches Coaching Programme: Evaluation report and executive summary*. London, UK: Education Endowment Foundation.
- *Maxwell, B., Connolly, P., Demack, S., O'Hare, L., Stevens, A., & Clague, L. (2014). *TextNow Transition Programme: Evaluation report and executive summary*. London, UK: Education Endowment Foundation.
- McLaughlin, M., & DeVoogd, G. (2018). Reading comprehension, critical understanding: Research-based practice. In D. Lapp & D. Fisher (Eds.), *Handbook of research on teaching the English language arts* (4th ed., pp. 85–110). New York, NY: Routledge.
- *Meisch, A., Hamilton, J., Chen, E., Quintanilla, P., Fong, P., Gray-Adams, K., ... Thornton, N. (2011). *Striving Readers study: Targeted and whole-school interventions—year 5*. Rockville, MD: Westat.
- *Merrell, C., & Kasim, A. (2015). *Butterfly Phonics: Evaluation report and executive summary*. London, UK: Educational Endowment Foundation.
- *Meyers, C.V., Molefe, A., Brandt, W.C., Zhu, B., & Dhillon, S. (2016). Impact results of the eMINTS professional development validation study. *Educational Evaluation and Policy Analysis*, 38(3), 455–476. <https://doi.org/10.3102/0162373716638446>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), article e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- National Center for Education Statistics. (2016). *The Nation's Report Card: 2015 mathematics and reading at grade 12*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016108>
- Newman, D.L., Kundert, D.K., Spaulding, D.T., White, S.P., & Gifford, T.A. (2012). *Striving Readers local evaluation*. Albany: University at Albany, State University of New York.
- Nomi, T. (2015). “Double-dose” English as a strategy for improving adolescent literacy: Total effect and mediated effect through classroom peer ability change. *Social Science Research*, 52, 716–739. <https://doi.org/10.1016/j.ssresearch.2014.12.012>
- OECD. (2013). *OECD skills outlook 2013: First results from the Survey of Adult Skills*. Paris, France: Author.
- OECD. (2016). *PISA 2015 results: Vol. I. Excellence and equity in education*. Paris, France: Author.
- *Olson, C.B., Kim, J.S., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D.A., ... Land, R.E. (2012). Enhancing the interpretive reading and analytical writing of mainstreamed English learners in secondary school: Results from a randomized field trial using a cognitive strategies approach. *American Educational Research Journal*, 49(2), 323–355. <https://doi.org/10.3102/0002831212439434>
- *Olson, C.B., Matuchniak, T., Chung, H.Q., Stumpf, R., & Farkas, G. (2016). Reducing achievement gaps in academic writing for Latinos and English learners in grades 7–12. *Journal of Educational Psychology*, 109(1), 7–12. <https://doi.org/10.1037/edu0000095>
- Oslund, E.L., Clemens, N.H., Simmons, D.C., & Simmons, L.E. (2018). The direct and indirect effects of word reading and vocabulary on adolescents' reading comprehension: Comparing struggling and adequate comprehenders. *Reading and Writing*, 31(2), 355–379. <https://doi.org/10.1007/s11145-017-9788-3>
- Paul, S.-A.S., & Clarke, P.J. (2016). A systematic review of reading interventions for secondary school students. *International Journal of Educational Research*, 79, 116–127. <https://doi.org/10.1016/j.ijer.2016.05.011>
- Pearson, P.D., & Cervetti, G. (2013). The psychology and pedagogy of reading processes. In W.M. Reynolds & G.E. Miller (Eds.), *Handbook of psychology: Vol. 7. Educational psychology* (2nd ed., pp. 257–281). Hoboken, NJ: John Wiley & Sons.
- Polanin, J.R., Tanner-Smith, E.E., & Hennessy, E.A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research*, 86(1), 207–236. <https://doi.org/10.3102/0034654315582067>
- Pressley, M. (2003). Psychology of literacy and literacy instruction. In W.M. Reynolds & G.E. Miller (Eds.), *Handbook of psychology: Vol. 7. Educational psychology* (pp. 333–356). Hoboken, NJ: John Wiley & Sons.

- Pyle, N., Vasquez, A.C., Lignugaris/Kraft, B., Gillam, S.L., Reutzel, D.R., & Olszewski, A., ... Pyle, D. (2017). Effects of expository text structure interventions on comprehension: A meta-analysis. *Reading Research Quarterly*, 52(4), 469–501. <https://doi.org/10.1002/rrq.179>
- Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Roseth, C.J., Johnson, D.W., & Johnson, R.T. (2008). Promoting early adolescents' achievement and peer relationships: The effects of cooperative, competitive, and individualistic goal structures. *Psychological Bulletin*, 134(2), 223–246. <https://doi.org/10.1037/0033-2909.134.2.223>
- *Rutt, S., Kettlewell, K., & Bernardenelli, D. (2015). *Catch Up* Literacy: Evaluation report and executive summary*. London, UK: Educational Endowment Foundation.
- Scammacca, N.K., Roberts, G.J., Cho, E., Williams, K.J., Roberts, G., Vaughn, S.R., & Carroll, M. (2016). A century of progress: Reading interventions for students in grades 4–12, 1914–2014. *Review of Educational Research*, 86(3), 756–800. <https://doi.org/10.3102/0034654316652942>
- Scammacca, N., Roberts, G., Vaughn, S., 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: Center on Instruction, RMC Research.
- *Schenck, A., Feighan, K., Coffey, D., & Rui, N. (2011). *Memphis Striving Readers project: Evaluation report, year 4*. Philadelphia, PA: Research for Better Schools; Arlington, VA: RMC Research.
- *Schenck, A., Jurich, S., Frye, M., Lammert, J., & Sayko, S. (with Najera, K., Taylor, L., & Willard, T.). (2012). *Evaluation report/impact study: Virginia Striving Readers Intervention Initiative (VSR II)*. Arlington, VA: RMC Research.
- *Schiller, E., Wei, X., Thayer, S., Blackorby, J., Javitz, H., & Williamson, C. (2012). *A randomized controlled trial of the impact of the Fusion Reading Intervention on reading achievement and motivation for adolescent struggling readers*. Evanston, IL: Society for Research on Educational Effectiveness.
- Shanahan, C., & Shanahan, T. (2018). Disciplinary literacy. In D. Lapp & D. Fisher (Eds.), *Handbook of research on teaching the English language arts* (4th ed., pp. 281–309). New York, NY: Routledge.
- Shannon, L., & Grant, B.-J. (2015). *A final report for the evaluation of the Achieve3000 programs*. Charlottesville, VA: Magnolia Consulting.
- *Shapley, K., Sheehan, D., Maloney, C., & Caranikas-Walker, F. (2009). *Evaluation of the Texas Technology Immersion Pilot: Final outcomes for a four-year study (2004–05 to 2007–08)*. Austin: Texas Center for Educational Research.
- *Sibieta, L. (2016). *REACH: Evaluation report and executive summary*. London, UK: Education Endowment Foundation.
- *Simmons, D., Fogarty, M., Oslund, E.L., Simmons, L., Hairrell, A., Davis, J., ... Fall, A.-M. (2014). Integrating content knowledge-building and student-regulated comprehension practices in secondary English language arts classes. *Journal of Research on Educational Effectiveness*, 7(4), 309–330. <https://doi.org/10.1080/19345747.2013.836766>
- *Simon, A.J., Tunik, J., Alemany, J., Zhu, J., Zacharia, J., Ramsey, L., ... Mendes, R. (2011). *Chicago Public Schools Striving Readers Initiative: Year four evaluation report*. New York, NY: Metis.
- *Sivin-Kachala, J., & Bialo, E.R. (2012). *Program evaluation research on Expert 21 in grades 6–8 for Scholastic Education*. New York, NY: Interactive Educational Systems Design.
- Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher*, 15(9), 5–11. <https://doi.org/10.3102/0013189X015009005>
- Slavin, R.E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.). Boston, MA: Allyn & Bacon.
- Slavin, R.E. (2015). Cooperative learning in schools. In J.D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd ed., pp. 881–886). Amsterdam, Netherlands: Elsevier.
- *Slavin, R.E., Chamberlain, A., Daniels, C., & Madden, N.A. (2009). *The Reading Edge: A randomized evaluation of a middle school cooperative reading program*. *Effective Education*, 1(1), 13–26. <https://doi.org/10.1080/19415530903043631>
- 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. <https://doi.org/10.1598/RRQ.43.3.4>
- 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), 1–26. <https://doi.org/10.1016/j.edurev.2010.07.002>
- *Somers, M.A., Corrin, W., Sepanik, S., Salinger, T., Levin, J., & Zmach, C. (with Wong, E.). (2010). *The Enhanced Reading Opportunities Study final report: The impact of supplemental literacy courses for struggling ninth-grade readers* (NCEE 2010-4021). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- *Sprague, K., Zaller, C., Kite, A., & Hussar, K. (2012). *Springfield-Chicopee School District Striving Readers (SR) Program: Final report years 1–5: Evaluation of implementation and impact*. Providence, RI: The Education Alliance, Brown University
- *Stevens, R.J., & Durkin, S. (1992). *Using student team reading and student team writing in middle schools: Two evaluations* (Report No. 36). Baltimore, MD: Center for Research on Effective Schooling for Disadvantaged Students, Johns Hopkins University.
- Stipek, D. (2002). *Motivation to learn: Integrating theory and practice* (4th ed.). Boston, MA: Allyn & Bacon.
- *Styles, B., Stevens, E., Bradshaw, S., & Clarkson, R. (2014). *Vocabulary Enrichment Intervention Programme: Evaluation report and executive summary*. London, UK: Education Endowment Foundation.
- *Swanlund, A., Dahlke, K., Tucker, N., Kleidon, B., Kregor, J., Davidson-Gibbs, D., & Hallberg, K. (2012). *Striving Readers: Impact study and project evaluation report: Wisconsin Department of Public Instruction (with Milwaukee Public Schools)*. Naperville, IL: American Institutes for Research.
- Uccelli, P., Phillips Galloway, E., Barr, C.D., Meneses, A., & Dobbs, C.L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly*, 50(3), 337–356. <https://doi.org/10.1002/rrq.104>
- U.S. Department of Education. (2017). *Reimagining the role of technology in education: 2017 national education technology plan update*. Washington, DC: Office of Educational Technology, U.S. Department of Education.
- *Vaden-Kiernan, M., Caverly, S., Bell, N., Sullivan, K., Fong, C., Atwood, E., & Jones, D. (2012). *Louisiana Striving Readers: Final evaluation report*. Austin, TX: Southwest Educational Development Laboratory.
- Vaughn, S., & Fletcher, J.M. (2012). Response to Intervention with secondary school students with reading difficulties. *Journal of Learning Disabilities*, 45(3), 244–256. <https://doi.org/10.1177/0022219412442157>
- *Vaughn, S., Klingner, J.K., Swanson, E.A., Boardman, A.G., Roberts, G., Mohammed, S.S., & Stillman-Spisak, S.J. (2011).

- Efficacy of collaborative strategic reading with middle school students. *American Educational Research Journal*, 48(4), 938–964. <https://doi.org/10.3102/0002831211410305>
- *Vaughn, S., Roberts, G., Klingner, J.K., Swanson, E.A., Boardman, A., Stillman-Spisak, S.J., Leroux, A.J. (2013). Collaborative strategic reading: Findings from experienced implementers. *Journal of Research on Educational Effectiveness*, 6(2), 137–163. <https://doi.org/10.1080/19345747.2012.741661>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. & Trans.). Cambridge, MA: Harvard University Press.
- Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review*, 36(4), 541–561.
- Wanzek, J., Vaughn, S., Scammacca, N.K., Metz, K., Murray, C.S., & Roberts, G. (2013). Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research*, 83(2), 163–195. <https://doi.org/10.3102/0034654313477212>
- Wexler, J., Vaughn, S., Edmonds, M., & Reutebuch, C.K. (2008). A synthesis of fluency interventions for secondary struggling readers. *Reading and Writing*, 21(4), 317–347. <https://doi.org/10.1007/s11145-007-9085-7>
- What Works Clearinghouse. (2017). *What Works Clearinghouse: Standards handbook version 4.0*. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf
- *Wijekumar, K., Meyer, B.J.F., & Lei, P. (2017). Web-based text structure strategy instruction improves seventh graders' content area reading comprehension. *Journal of Educational Psychology*, 109(6), 741–760. <https://doi.org/10.1037/edu0000168>
- Wright, T., & Cervetti, G. (2017). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, 52(2), 203–226. <https://doi.org/10.1002/rrq.163>
- Zief, S.G., Lauver, S., & Maynard, R.A. (2006). *Impacts of after-school programs on student outcomes*. *Campbell Systematic Reviews*, 2(3). <https://doi.org/10.4073/csr.2006.3>

Submitted October 26, 2017
 Final revision received July 25, 2018
 Accepted July 26, 2018

ARIANE BAYE is an assistant professor in the Department of Education and Training at the University of Liege, Belgium, and the director of the Center for Research and Interventions on Dropout and Exclusion; email ariane.baye@ulg.ac.be. Her research focuses on effective interventions to prevent school failure and dropout.

AMANDA INNS is an associate research scientist at the Center for Research and Reform in Education at Johns Hopkins University, Baltimore, Maryland, USA; email ainns1@jhu.edu. Her research interests include school-based health models, program evaluation, and systematic reviews.

CYNTHIA LAKE is a senior research associate at the Center for Research and Reform in Education at Johns Hopkins University, Baltimore, Maryland, USA; email clake1@jhu.edu. Her research interests include comprehensive school reform, school improvement, and educational research methods and statistics.

ROBERT E. SLAVIN (corresponding author) is the director of the Center for Research and Reform in Education at Johns Hopkins University, Baltimore, Maryland, USA; email rsalvin@jhu.edu. He has authored or coauthored more than 300 articles and book chapters on such topics as cooperative learning, comprehensive school reform, research reviews, evidence-based reform, and vision and learning.

APPENDIX

Program Descriptions

Tutoring Interventions

One-to-One Tutoring

Catch Up Literacy is a structured one-to-one tutoring intervention. Paraprofessionals provide 15-minute sessions to struggling readers twice a week over the course of a school year. Each session includes prepared reading, reading aloud, discussing the text, and linked writing. A study of Catch Up Literacy in year 6 (Rutt et al., 2015) found a marginally statistically significant difference favoring the tutored students ($ES = 0.16$, $p = .08$).

The Perry Beeches Coaching Programme provides struggling readers with one hour of one-to-one tutoring every two weeks. Coaches tailor activities according to

students' needs. A study of this program by Lord et al. (2015) found a large positive effect ($ES = 0.36$, $p < .01$).

REACH Tutoring provides struggling readers with one-to-one tutoring in 35-minute sessions once a week for 20 weeks. The tutors are specially trained paraprofessionals. Sibieta (2016) evaluated two very similar variations and found a mean effect size of 0.42.

One-to-Small Group Tutoring

Butterfly Phonics uses formal phonics instruction, understanding the global aspects of a text, and class discussion of text meaning to improve reading comprehension. The program is delivered to groups of six to eight students by a trained practitioner and an assistant. Merrell and Kasim

(2015) evaluated Butterfly Phonics and found an effect size of 0.30 ($p < .001$).

Rapid Phonics combined with Sound Discovery is a small-group tutoring program designed to improve decoding skills and reading fluency using structured instruction in letter–sound correspondence. Students in the experimental group were taught by specialists in groups of up to four taken out of their regular classes, whereas control students continued their schooling as usual. In the evaluation (King & Kasim, 2015), the effect size was statistically nonsignificant ($ES = -0.05$, n.s.). The failure to find positive effects, unique among studies of nonvolunteer tutors, may have been due to a focus on phonics, when the outcome test (New Group Reading Test) has a clear focus on comprehension. Also at 12 weeks, this study provided significantly less tutoring than did other programs.

Volunteer Tutoring

TextNow Transition Programme utilizes volunteer coaches to provide daily 20-minute sessions focused on encouraging students to read for pleasure. A randomized trial of the program in the United Kingdom (Maxwell et al., 2014) found no statistically significant effects on reading comprehension ($ES = -0.06$, n.s.).

Cross-Age Tutoring

Paired Reading is a cross-age tutoring approach with the goal of improving general literacy. Year 9 students work with year 7 students to choose, read, and discuss a text. Paired Reading showed no statistically significant effects ($ES = -0.02$, n.s.) in a randomized experiment (Lloyd et al., 2015).

Cooperative Learning Approaches

The Reading Edge, adapted from a program called Student Team Reading, is a cooperative learning program for middle schools in which students work in four- or five-member teams to help one another build reading skills. Students engage in partner reading, story retelling, story-related writing, word mastery, and story structure activities to prepare themselves and their teammates for individual assessments that form the basis for team scores. Instruction focuses on explicit teaching of metacognitive strategies. Across three studies of The Reading Edge, the weighted mean effect size was 0.15 (Slavin et al., 2009; Stevens & Durkin, 1992). The outcomes in the Slavin et al.'s study were statistically significantly positive ($ES = 0.15$, $p < .05$).

Collaborative Strategic Reading teaches reading comprehension strategies to students working in small cooperative learning groups. The weighted mean effect size across three studies was 0.05 (Denver Public Schools, 2016; Vaughn et al., 2011, 2013).

Whole-School Approaches

Whole-School Approaches With Organizational Elements

BARR is a whole-school reform approach focused on developmental, academic, and structural challenges during ninth grade. BARR is used in all major subjects to attempt to increase student achievement by improving students' socioemotional skills, building positive student–teacher relationships, and solving nonacademic barriers to learning, such as truancy and behavioral problems. The strategy focuses on building students' personal assets and reducing substance abuse, delinquency, and other problems. BARR staff closely monitor student achievement, including real-time analysis of student data. Students take English, math, and science or social studies in a block to build connections among students and teachers. Teachers in each block meet regularly to review the progress of at-risk students. Extensive PD and coaching are provided to teachers and school leaders. Corsello and Sharma (2015) found a positive effect of BARR on reading ($ES = 0.14$, $p < .01$), and a larger study by T. Borman et al. (2017) also found statistically significant positive effects ($ES = 0.08$, $p < .05$). The weighted mean across the two studies was 0.09.

Talent Development High School is a whole-school reform model for ninth graders. Within the program, Strategic Reading and Student Team Literature is the reading component. Students receive a double dose of reading and math, amounting to 90 minutes a day for each subject. In Strategic Reading and Student Team Literature, students work in small, interdependent cooperative learning groups using structured partner discussion guides that provide background and vocabulary, and comprehension questions provide minilessons on specific comprehension strategies.

Balfanz et al. (2004) carried out an evaluation of Talent Development High School in high-poverty high schools in Baltimore, Maryland, comparing against control schools that also provided double-dose reading. There was a statistically nonsignificant effect size of 0.32 (significance was not attained because there were only six schools).

Whole-School Approaches Without Organizational Elements

Teacher Effectiveness Enhancement Programme is a U.K. PD approach for secondary teachers involving all major

subjects. It includes formative assessment, thinking skills, cooperative learning, and effective use of technology. An evaluation by the Institute for Effective Education (2016) found no positive effects ($ES = -0.04$, n.s.).

Chicago Striving Readers provides students with technology tools (media and listening centers and handheld computers) to support their learning. All teachers are encouraged to assign partner reading in social studies, science, and math. A multiyear evaluation found no statistically significant effects (Simon et al., 2011; $ES = -0.01$, n.s.).

Project CRISS is a PD approach designed to help teachers of all subjects use proven reading comprehension strategies. The project provides summer institutes to local facilitators, who then create local teacher-to-teacher study groups. An evaluation of Project CRISS (Kushman, Hanita, & Raphael, 2011) found no statistically significant effects ($ES = 0.05$, n.s.).

eMINTS is a schoolwide program that provides extensive PD to teachers to help them with technology integration, inquiry-based learning, and high-quality lesson design. In an evaluation by Meyers, Molefe, Brandt, Zhu, and Dhillon (2016), the mean reading effect size for eMINTS was -0.06 (n.s.).

Texas Technology Immersion Pilot was a three-year evaluation of a technology immersion intervention covering language arts, math, science, and social studies in grades 6–8. Contractors provided schools with wireless, mobile computing devices for every student and teacher; productivity, communication, and presentation software; online resources supporting state standards; online assessments linked to state standards; extensive PD; and initial and ongoing technical support. A study by Shapley, Sheehan, Maloney, and Caranikas-Walker (2009) found a mean effect size of 0.06 (n.s.).

Writing-Focused Approaches

Pathway is a PD program used primarily with mainstreamed Latino ELs. Teachers learn how to teach cognitive strategies and process writing. A study by Olson et al. (2012) found statistically significant differences ($ES = 0.07$, $p < .05$), and the weighted average across the two Pathway studies (Olson et al., 2012, 2016) was 0.08 .

Expository Reading and Writing Course is a program for 12th graders designed to prepare them to pass the California Early Placement Test, used in the California State University system to determine whether freshmen must take noncredit remedial English courses. The emphasis of the program is on discussion of text meaning, developing critical thinking skills, encouraging group discussions, and developing writing skills in multiple genres. Fong et al. (2015) found

statistically significant positive effects of the program ($ES = 0.13$, $p < .05$).

Content-Focused Approaches

Reading Apprenticeship is a family of programs designed to improve reading comprehension by integrating metacognitive strategy instruction into content areas, such as science and social studies. The model incorporates extensive reading, teaching of comprehension skills, and collaborative sensemaking. A study of the core model by Greenleaf et al. (2011) found no statistically significant outcomes ($ES = 0.03$, n.s.).

Reading Apprenticeship Improving Secondary Education is a whole-class version of Reading Apprenticeship. A study of this version by Fancsali et al. (2015) found statistically nonsignificant positive effects ($ES = 0.14$, n.s.).

Internet-based Reading Apprenticeship Improving Science Education is a form of Reading Apprenticeship that provides PD online. Jaciw, Schellinger, Lin, Zacamy, and Toby (2016) found no positive effects of this approach ($ES = 0.00$, n.s.).

RAAL is an adaptation of Reading Apprenticeship designed for struggling readers and taught during an extra daily reading period. A study of RAAL by Somers et al. (2010) found statistically significant positive effects ($ES = 0.10$, $p < .05$). Combining RAAL with the three other variations of Reading Apprenticeship produced a weighted mean of 0.07 . Because of the RAAL study, the whole Reading Apprenticeship family of studies was considered to meet ESSA standards for “strong” evidence.

Every Classroom, Every Day is a structured literacy curriculum based on authentic expository texts. Early et al. (2016) found no statistically significant effects of the program on reading ($ES = 0.06$, n.s.).

Intelligent Tutoring System for the Structure Strategy is a web-based approach in which students are taught to comprehend nonfiction text by categorizing text structures using key elements in the text to find the main idea, activating prior knowledge, supporting cognitive monitoring, and using graphic organizers and flowcharts to summarize texts. This approach is used during regular language arts classes with the assistance of a paraprofessional. Animated tutors model and guide learners. Students practice, take regular assessments, and proceed at their own pace through self-instructional units. A study by Wijekumar et al. (2017) found a statistically significant positive effect for seventh graders ($ES = 0.18$, $p < .05$).

Content Knowledge-Building and Student-Regulated Comprehension Practices trains teachers to introduce texts with critical questions. Then, students work in pairs to analyze the text and answer critical

questions. In a study by Simmons et al. (2014), the effect size was -0.01 (n.s.).

Vocabulary-Focused Approaches

Academic Language Instruction for All Students is a vocabulary intervention designed to be used in regular English language arts classrooms including many language-minority students. Each cycle of lessons is based on one informational text, from which are extracted a small number of high-utility and abstract words. The intervention includes a variety of whole-group, small-group, and independent activities. The weighted mean effect size across two studies was 0.07 (n.s.; Lesaux et al., 2010, 2014).

Word Generation is an approach to vocabulary building in which students are encouraged to discuss and read about topics containing target words believed to be important but not already in students' speaking or reading vocabularies. Lawrence, Francis, Paré-Blagoev, and Snow (2017) evaluated reading effects of Word Generation and found no statistically significant differences on reading ($ES = 0.05$, n.s.).

Vocabulary Enrichment Intervention Programme combines vocabulary teaching, phonics, and connected text. Teachers are given the flexibility to use the components as needed by their students. A U.K. evaluation of the program by Styles, Stevens, Bradshaw, and Clarkson (2014) found no statistically significant outcomes ($ES = 0.06$, n.s.).

Strategy-Focused Instruction

SIM: Xtreme Reading

SIM: Xtreme Reading is the main version of the SIM designed for struggling readers. Students reading two to five years below grade level are usually given an additional daily reading period, but in one of four studies (Faddis et al., 2011), Xtreme Reading was taught during an extra reading period in high schools but not in middle schools. Across the four qualifying studies (Faddis et al., 2011; Somers et al., 2010; Sprague et al., 2012), the mean effect size for Xtreme Reading was 0.09 ($p < .01$).

SIM: Content Literacy Curriculum

A large, two-year study by Corrin et al. (2012) found statistically nonsignificant positive effects on reading ($ES = 0.09$, n.s.).

SIM: Learning Strategies Curriculum

A study by Cantrell, Almasi, Rintamaa, and Carter (2016) found statistically significant positive effects on reading ($ES = 0.10$, $p < .05$).

SIM: Fusion Reading

A study by Schiller et al. (2012) found statistically nonsignificant positive effects ($ES = 0.07$, n.s.).

SIM: Adolescent Literacy Model

SIM: Adolescent Literacy Model was evaluated by Cantrell, Almasi, Carter, and Rintamaa (2011) in Kentucky middle and high schools. The researchers reported a statistically nonsignificant effect size of 0.10 .

Personalization Approaches

Thinking Reader teaches reading comprehension skills to struggling readers. It provides students novels with a range of difficulty. Animated coaches and peers on the computer model comprehension strategies (e.g., summarizing, questioning, predicting, visualizing) and prompt students to use them. In a study by Drummond et al. (2011), the average effect size was 0.01 (n.s.).

Schoolwide Enrichment Model-Reading exposes students to a variety of books. Students independently read self-selected, challenging books and meet their teacher individually a few minutes every one to two weeks to discuss reading strategies and respond to higher level questions. A study by Little, McCoach, and Reis (2014) found statistically nonsignificant positive effects ($ES = 0.10$, n.s.).

Achieve3000 is an online literacy program that provides nonfiction reading content and teaches metacognitive skills to improve the comprehension of informational texts. In an evaluation by Shannon and Grant (2015), the effect size was 0.29 across grades 6 and 9 ($p < .05$).

SuccessMaker is an adaptive K-8 computer-based reading program. It provides individualized reading activities, game-like environments, interactive aids, and a reporting system to inform teachers of student progress. A study by Gatti (2011) found a statistically nonsignificant effect size for seventh graders of 0.11 (n.s.).

Accelerated Reader is a widely used U.S. program, but the only qualifying evaluation in secondary reading took place in the United Kingdom (Gorard et al., 2015). The program provides students with a wide range of books at their reading level, as determined by an online test. Online comprehension tests are provided for each book, and students can earn points based on completing many books at a high readability level. Gorard et al.'s evaluation found a statistically significant effect size of 0.24 ($p < .05$).

iLit is a digital instruction approach for struggling readers. Students choose among more than 500 e-books and work on vocabulary and comprehension strategies. Students work independently, keeping online journals,

answering questions, and discussing books in groups. A two-year study of iLit by Gatti (2016) found an effect size of 0.09 (n.s.).

Group/Personalization Rotation Approaches

READ 180 is a supplemental model used 90 minutes each day with struggling readers. Students rotate through whole-group instruction, independent reading, small-group direct instruction, and use of READ 180 adaptive software. Across all five qualifying studies of READ 180 (Lang et al., 2009; Meisch et al., 2011; Schenck et al., 2011; Sprague et al., 2012; Swanlund et al., 2012), the mean effect size was 0.09. Outcomes were identical for studies that did or did not provide an additional reading period.

Expert 21 was designed for students who have graduated from READ 180. Expert 21 provides student texts and supportive online materials focused on building language arts, writing, and comprehension skills and includes whole-class and small group-discussions, teaching of metacognitive skills such as graphic organizers, and collaborative projects. Sivin-Kachala and Bialo (2012) found statistically nonsignificant positive effects (ES = 0.15, n.s.).

System 44 is a version of READ 180 for adolescent readers who have not mastered basic phonics and decoding skills. The program focuses on decoding, fluency, and comprehension. In daily 60-minute lessons, the teacher gives five to 10 minutes of whole-class instruction, the students spend 25 to 30 minutes working in small groups or individually, and they then receive 20 to 25 minutes of computer-delivered instruction. Like READ 180, System 44 is always used during supplemental reading time, but in a study by Beam, Faddis, and Hahn (2011), the control group also had additional reading time, so there were no time differences. Outcomes were statistically significantly negative on the Test of Silent Reading Efficiency and Comprehension (ES = -0.24 , $p < .05$) and statistically nonsignificantly negative on the California Standards Test (ES = -0.04), for a mean effect size of -0.14 . A study that provided supplemental reading time (Beam & Faddis, 2012) found a statistically significant positive impact (ES = 0.20, $p < .05$). Across the two studies, the mean effect size was 0.03.

Passport Reading Journeys is a supplemental literacy curriculum that provides 15 two-week sequences of lessons mixing whole-class and small-group lessons, as well as individualized computer-based practice. The curriculum focuses on reading comprehension strategies, vocabulary, word study, and writing, using mainly science and social studies topics. Across three studies (Dimitrov et al., 2012; Schenck, Jurich, Frye, Lammert, & Sayko, 2012; Vaden-Kiernan et al., 2012), the weighted mean

effect size was 0.07. Vaden-Kiernan et al.'s study found statistically significant positive effects on the Group Reading Assessment and Diagnostic Evaluation (ES = 0.27, $p < .05$) but not the Integrated Louisiana Educational Assessment Program (ES = -0.01), for a mean of 0.12.

Comprehension Circuit Training uses content delivered on tablet computers to teach reading comprehension skills. Following video instruction, students work with a partner to practice lesson content. Students cycle through four major components, focusing on vocabulary skills, prereading, reading of eBooks, and comprehension quizzes. Across two studies (Fogarty et al., 2014, 2017), the mean effect size was 0.13. In Fogarty et al.'s (2017) study, statistically significant positive effects were found on the Test of Silent Reading Efficiency and Comprehension (ES = 0.24, $p < .05$) but not other measures.

Prentice Hall Literature combines offline textbooks with online components. Online material includes vocabulary games, audios, and videos. A study by Eddy, Ruitman, Hankel, and Sloper (2010) found statistically nonsignificant negative impacts (ES = -0.10 , n.s.).

Strategic Adolescent Reading Intervention uses core novels and other engaging texts within thematic units, fluency passages, comprehension instruction, reciprocal teaching strategies, and student discussion and debate. Kim et al. (2017) conducted a randomized trial of with low-achieving middle school students, and although the overall effects of the program were statistically nonsignificantly positive (ES = 0.15, $p =$ n.s.), the program demonstrated statistically significant positive effects on the Efficiency of Basic Reading subtest of the Reading Inventory and Scholastic Evaluation (ES = 0.21, $p < .05$).

Read to Achieve provides lessons on content area and narrative texts and incorporates small-group collaboration and independent activities. A study with low-achieving middle school students (Deussen, Scott, Nelsestuen, Roccograndi, & Davis, 2012) found statistically nonsignificant positive effects on reading (ES = 0.10, n.s.).

Reading Intervention through Strategy Enhancement relies on teachers' capacity to build effective curriculum for struggling readers. During classes, students are given the opportunity to read independently, work in small groups, and receive whole-group lessons. A study by Lang et al. (2009) found positive effects on reading for moderate-risk students (ES = 0.27, $p < .04$) but not high-risk students (ES = -0.06 , n.s.). The weighted mean effect size was 0.16 ($p < .05$).

Intensive Group Approaches

Intensive group approaches are programs other than tutoring designed to catch students up on decoding

and word study skills, as well as basic comprehension strategies, which other students would have mastered in elementary school. The studies summarized in Table 12 had a weighted mean effect size of 0.00 (n.s.).

Reading Excellence: Word Attack and Rate Development Strategies is a one-year supplemental intervention given by trained teachers five times a week to develop reading skills in low achievers. An evaluation by Newman, Kundert, Spaulding, White, and Gifford (2012) found a mean effect size of 0.09 (n.s.).

Kentucky Cognitive Literacy Model is an intervention for struggling readers focused on teaching comprehension strategies, vocabulary, study skills, and writing. A study of the model by Cantrell, Carter, and Rintamaa (2012) found no statistically significant outcomes ($ES = -0.06$).

REACH provides explicit, intensive instruction to struggling students in grades 6–12. It focuses on phonics, fluency, word knowledge, and reasoning skills; narrative structure and writing skills; and spelling. Lang et al. (2009) evaluated REACH and found a mean effect size of -0.02 (n.s.).



Explore the Latest Topics and Trends in Literacy Education With ILA

Interested in ILA's views on current topics and trends in literacy education?

INTERNATIONAL
LITERACY
ASSOCIATION

Get our research-based perspectives on assessment, diversity, early literacy, education policy, equity in literacy education, teacher preparation, and more at literacyworldwide.org/statements.