Reading Programmes for Secondary Students: Evidence Review

July 2019

Ariane Baye
(University of Liege, Belgium)
Robert E. Slavin
Cynthia Lake
Amanda Inns
(Johns Hopkins University, USA)
Jonathan Haslam
(Institute for Effective Education, England)

Contents

Abstract

Recent initiatives in the UK and the US have added greatly to the amount and quality of research on the effectiveness of secondary reading programmes, especially programmes for struggling readers. This review of the experimental research on secondary reading programmes focuses on 69 studies that used random assignment (n=62) or high-quality quasi-experiments (n=7) to evaluate outcomes of 51 programmes on widely accepted measures of reading. Categories of programmes using one-to-one and small-group tutoring, cooperative learning, whole-school approaches including organisational reforms such as teacher teams, and writing-focused approaches showed positive outcomes. Individual approaches in a few other categories also showed positive impacts. These include programmes emphasising arts, humanities and science, structured strategies, and personalised and group/personalisation rotation approaches for struggling readers. Programmes that provide a daily extra period of reading and those utilising technology were no more effective, on average, than programmes that did not provide these resources. The findings suggest that secondary readers benefit more from socially and cognitively engaging teaching than from additional reading periods or technology.

Key words: secondary reading programmes, research reviews, best-evidence synthesis, middle school reading, high school reading

A Synthesis of Quantitative Research on Reading Programmes for Secondary Students

The reading performance of students in England's secondary schools has been static in recent years. In 2015, the Programme for International Student Assessment (PISA) showed that there had been no significant change in the average reading scores of England's 15-year-olds since 2006. There were nine countries where the mean reading score was at least four months ahead of England, although we might console ourselves with the knowledge that there were 41 countries where the mean reading score was at least four months behind ours.

The top 10 percent of English pupils perform strongly in reading, in comparison with other countries, with relatively few countries performing significantly better. The gap between the best and worst performing readers is similar to most countries.

There are bigger differences in achievement amongst 15-year-olds who attend the same school than there are differences in achievement between pupils who attend different schools. This is a common finding across countries with a comprehensive education system.

Also shared with other countries is a difference in the performance of boys and girls. Boys on average are nine months of schooling behind girls when it comes to reading.

Another OECD study (http://www.oecd.org/skills/piaac/Country%20note%20-%20United%20Kingdom.pdf) found that in England, adults aged 55-65 performed better than 16-24 year-olds in both literacy and numeracy. In fact, England was the only country (of the 22 that took part) where the oldest age group had higher proficiency in both literacy and numeracy than the youngest age group (after other factors were taken into account).

Ensuring that students are confident readers on entering secondary school is vital to their long-term success. Less than one fifth of pupils who did not reach Level 4 in English overall in 2008

(the then-expected standard) went on to achieve a C or above at English GCSE, compared with four-fifths of pupils who achieved Level 4 or above. Only 11% of pupils who did not reach Level 4 in English overall in 2008 went on to achieve five or more A*-C grades at GCSE, including English and maths, compared to 72% of pupils who achieve a Level 4 or above. (https://educationendowmentfoundation.org.uk/public/files/Publications/EEF_Publications_Evid enceBrief_ReadingAtTheTransition.pdf)

The Rising Importance of Evidence

There is a new movement toward the use of evidence in education, especially in the UK and the US. This movement is intended to identify and disseminate proven practices and interventions, thereby improving outcomes for students. In 2012, the Education Endowment Foundation launched a £10 million fund that tested 24 pilot projects to improve outcomes, particularly in literacy, for pupils during the transition from primary to secondary school. (https://educationendowmentfoundation.org.uk/news/eef-transition-catch-up-round-open/) The U.S. Department of Education's Striving Readers programme (Boulay, Goodson, Frye, Blocklin, & Price, 2015) focused on secondary reading, and funding from Investing in Innovation (i3) and from the Institute for Education Sciences (IES) have also often gone to research on secondary reading (Herrera, Truckenmiller, & Foorman, 2016). As a result of these and other investments, the numbers of studies of secondary reading using rigorous research methods, especially cluster randomised designs with large samples, has increased dramatically.

Current Issues in Secondary Reading

Reading occupies a special place in secondary schools (Kamil et al., 2008). On one hand, the importance of reading is obvious, as success in all content areas depends on skilful and deep understanding of all sorts of texts. Accountability in the secondary years depends substantially on students' reading performance. Yet most secondary students do not have a separate reading class, so reading is at the same time the responsibility of all staff yet not the primary responsibility of any particular staff member.

Some secondary schools have introduced specific reading initiatives, such as DEAR (Drop Everything And Read) sessions to try and raise reading levels. (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/409409/Reading_the_next_steps.pdf)

The remarkable growth of technology has also redefined what counts as "literacy." Students must not only be conversant 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. 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.

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

Over the past ten years, several reviews of research on secondary reading programmes have been published, and these provide 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, the current review found 69

experiments (62 of which were randomised) 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 programmes for grades 4-12 (Years 5-13), and Edmonds et al. (2009) located 17. A review by Flynn, Zheng, & Swanson (2012) reported on only 10 studies of struggling readers in grades 4-8 (Years 5-9). A UK review by Cockroft & Atkinson (2017) also reviewed some of this literature.

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.

Some previous reviews examined secondary reading interventions and identified approaches associated with particularly positive reading outcomes. These include Slavin, Cheung, Groff, & Lake (2008), who concluded that secondary reading programmes that incorporated cooperative learning and other innovations in classroom teaching practices had the strongest effects on reading achievement in Years 7-13. Herrera et al. (2016) also reported that cooperative learning approaches and other methods providing extensive professional development were particularly likely to have positive outcomes. Dietrichson et al. (2017) found the strongest support for tutoring, feedback/progress monitoring, and cooperative learning.

Focus of the Review

The present review, adapted from an article by Baye, Lake, Inns, & Slavin (2019), synthesises research on reading outcomes of programmes designed for middle and high school students. It uses best evidence synthesis (Slavin, 1986), a method adapted from meta-analysis (see Cooper, 1998; Lipsey & Wilson, 2001) that includes narrative as well as numeric summaries of

8

the methods and findings of all studies meeting a common set of inclusion criteria. The review describes methods and outcomes of individual studies and programmes, and places 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 randomised, quantitative research evaluating secondary reading programmes. There is much other research on secondary reading, including qualitative and correlational methods and outcomes other than standardised 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 programmes and programme types are able to accelerate the reading achievement of secondary students and that is what we have attempted to provide.

Methods

Criteria for Inclusion

The review focused on a set of studies that met rigorous inclusion criteria. The criteria were designed to minimise bias and maximise potential replicability in schools not involved in the research. These were as follows.

1. Studies evaluated reading programmes for secondary schools, Years 7-13.

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 programme to 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 programmes took place in the UK or the US.

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 post-test 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 standardised tests were used, "total reading" or "total comprehension" 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 standardised tests for the secondary years.

9. Assessments made by developers or researchers were excluded, as such measures have been found to greatly overstate programme 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 programmes 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, due to the significant advances in technology since that date.

Literature Search Procedures

A broad literature search was carried out in an attempt to locate every study that could possibly meet the inclusion requirements. Electronic searches were made of educational databases (JSTOR, ERIC, EBSCO, Psych INFO, Dissertation Abstracts) using different combinations of key words (for example, "secondary students," "reading," "achievement") and the years 1990-2017. Results were then narrowed by subject area (for example, "reading intervention," "educational software," "academic achievement," "instructional strategies"). In addition to looking for studies by key terms and subject area, we conducted searches by programme name. Web-based repositories and education publishers' websites were also examined. These efforts were made 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 (2016) and ones reported online by i3, IES, EEF, and other funders and researchers. We contacted producers and developers of reading programmes to check whether they knew of studies that we had missed. Citations from previous reviews of secondary reading programmes or potentially related topics such as technology were further investigated. We also conducted searches of recent tables of contents of key journals from 2003 to 2017: *American Educational Research Journal, Reading Research Quarterly, Journal of Educational Research, Journal of Adolescent & Adult Literacy, Journal of Educational Psychology, British Journal of Educational Psychology, British Journal of Educational Research, International Journal of Educational Research,* and *Reading and Writing Quarterly.* Citations of

studies appearing in the studies found in the first wave were also followed up. This process was continued until no new studies were being found.

Effect Sizes

Effect sizes were computed as the difference between experimental and control individual student post-tests after adjustment for pretests and other covariates, divided by the unadjusted post-test control group standard deviation (SD). If the control group SD was not available, a pooled SD was used. Procedures described by Lipsey & Wilson (2001) were used to estimate effect sizes when unadjusted standard deviations were not available. If pretest and post-test means and SDs were presented but adjusted means were not, effect sizes for pretests were subtracted from effect sizes for post-tests.

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 (ANCOVA) controlling for pretests and, in most studies, demographic variables (e.g., ethnicity, free school meals).

When studies randomly assigned classes or schools to treatments or when they compared matched classes or schools, they should have used multilevel modelling such as Hierarchical Linear Modelling (HLM; Raudenbush & Bryk, 2002) to analyse 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 HLM, using a formula provided by the What Works

12

Clearinghouse (2017). In several studies (e.g., Balfanz et al., 2004; Stevens & Durkin, 1992a) large effect sizes were reported, but there were small numbers of experimental and control schools or classes. Accounting for clustering made these large-appearing effects non-significant. A programme is considered effective if it has at least one statistically significant positive effect, and no significant negative effects.

Statistical Procedures

Mean effect sizes across studies were calculated after assigning each study a weight based on inverse variance (Lipsey & Wilson, 2001), adjusted as suggested by Hedges (2007) which inflates the variances from school- and class-assigned studies. In combining across studies, we used a random-effects models as recommended by Borenstein et al. (2009) when there is a belief that there is no single "true" effect size, but a range of effect sizes that may depend on other factors. Weighted mean effect sizes and meta-analytic tests such as Q statistics were calculated in R (R Core Team, 2016) using the *metafor* package (Viechtbauer, 2010).

Programme Categories

Once inclusion criteria were met, the programmes studies evaluated were each placed in categories according to the most important and distinctive components, the key elements of a programme that distinguish it from other programmes. Category assignments were based on independent close reading of articles and websites by the authors, then debated among authors and resolved by consensus.

Research and theory supporting main programme components. The identified programme components and resulting categories were guided by two main sources. The first,

Reading Next (Biancarosa & Snow, 2006), identified fifteen elements of adolescent literacy programmes that support increased achievement. The second was the IES Practice Guide on Adolescent Literacy (Kamil et al., 2008). The recommendations of these two reports were combined and adapted to produce ten categories of programmes, each with distinguishing components that are 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 to help them comprehend what they read. These include clarification, summarisation, graphic organisers, outlining, and prediction, among others. 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. That 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; Kamil et al., 2008; Pressley, 2003).

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

The importance of professional development. All of the accepted studies provided significant professional development to teachers. Professional development is an essential element of school improvement (see Timperley et al., 2007). Some studies explained the amount and

14

quality of professional development provided to teachers but most did not, so we were unable to use the amount of professional development as a cross-cutting factor.

Programme categories are discussed in the following sections.

1.   **Tutoring**. In tutoring programmes, struggling readers receive one-to-one or small-group tutoring, in groups of one to four. Tutors can be teachers, teaching assistants, 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 teaching assistants have been very effective in primary schools (Slavin, Lake, Davis, & Madden, 2011; Edmonds et al., 2009; Wanzek & Vaughn, 2007), with some evidence supporting its effectiveness in secondary school (Jun, Ramirez, & Cumming, 2010).

Tutoring emphasises personalisation 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..

2.   **Cooperative learning programmes.** Cooperative learning programmes involve students working daily in small mixed-ability groups. Usually, cooperative learning groups have 4-5 members. The students are encouraged to help each other learn academic content, especially helping each other to learn and apply metacognitive comprehension strategies. Cooperative learning approaches have been shown to be among the most effective strategies for improving adolescent literacy (Slavin, Cheung, Groff, & Lake, 2008; Dietrichson et al., 2017; Herrera et al., 2016).

Cooperative learning may improve reading comprehension in several ways. First, it emphasises motivation through engagement with peers and encouragement from them, learning by explaining to peers and receiving explanations from them, and personalisation through

individualised feedback from peers and teachers (Slavin, 2015; Roseth, Johnson, & Johnson, 2008). Cooperative learning can also provide opportunities for participation in high-quality discussions of text that support and increase comprehension (Kamil et al., 2008).

3.   **Whole-school approaches**. Several programmes provide professional development for teachers of all major subjects, in a coordinated schoolwide plan to increase student motivation and achievement. These programmes may also build cross-disciplinary teaching teams and focus on social emotional skills. Previous research on comprehensive school reforms has demonstrated that some are effective when well implemented (Borman, Hewes, Overman, & Brown, 2003).   In secondary reading, providing all teachers with professional development in reading comprehension strategies is expected to give teachers across disciplines a common language and toolkit of effective strategies, giving students consistent comprehension instruction in many of their classes.

4.   **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, for example, Graham & Hebert, 2011; Graham et al., 2017). Writing about text may help improve comprehension (Fitzgerald & Shanahan, 2000). Classroom activities focused on writing instruction, such as learning about the writing process or specific writing skills instruction, 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 (Nelson & Calfee, 1998). 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).

16

5.   **Content-focused approaches**. In secondary schools, metacognitive strategies may be tailored to informational text, especially in the humanities or science. This reading instruction may be provided by English, reading, humanities, 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 that 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 specialised demands and require tailored comprehension skills (Lee & Spratley, 2010). This leads to students learning how to "read like a scientist" and "read like a historian", depending on the requirements of the text (Goldman, 2012).

6.   **Vocabulary-focused approaches**. Several programmes focus on developing students' vocabulary, expecting that vocabulary will generalise to building students' comprehension. These programmes 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 relationship with reading proficiency (Oslund, Clemens, Simmons, & Simmons, 2017; Uccelli, Galloway, Barr, Meneses, & Dobbs, 2015).

7.   **Strategy-focused instruction.** The Strategic Instruction Model (SIM) is a family of programmes all of which emphasise teaching students step-by-step approaches to cope with comprehension difficulties, decoding, writing, and other objectives. These metacognitive strategies include summarising, questioning, identifying the main idea, and using graphic organisers. Students learn mnemonics to recall how to accomplish key metacognitive objectives (Deshler & Schumaker, 2005). As noted earlier, teaching students these comprehension strategies has a positive impact on reading achievement.

17

8.   Personalisation approaches. Personalisation refers to instructional approaches that adapt to the learning levels, interests, or other characteristics of individual students. For example, Accelerated Reader and iLit both provide students with choices among hundreds of books at their own instructional level, so the software is personalised both for reading level and for student interests.

Today, personalised approaches usually use computers, but this is not a defining characteristic. For example, Accelerated Reader existed long before ebooks were made available.

The rationale for personalisation 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, personalised content may avoid frustrating students with work 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 (Stipek, 2002).

9.   Group/Personalisation Rotation Approaches. Several programmes rotate students through activities. Typically, one activity is teacher-directed (e.g., a traditional lesson), and one is personalised (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, while computers may be ideal for providing personalised practice and cooperative learning may be ideal for peer tutoring or projects (U.S. Department of Education, 2017).

10.   Intensive approaches. While 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. Such programmes also focus on comprehension, vocabulary, and other objectives, but

18

they are distinctively focused on identifying and remediating serious gaps in students' prior learning (Vaughn & Fletcher, 2012).

Cross-cutting factors are discussed in the following sections.

1.   **Extra class periods for reading instruction.** In addition to the ten strategy types listed above, two important cross-cutting factors were also analysed: extra daily periods for reading instruction and use of technology. Many of the qualifying programmes were provided to students in daily class periods in addition to ordinary reading or English periods, replacing music, art, or other ordinary parts of students' days. Such "double dose" strategies for reading have been popular as a response to accountability pressure to improve outcomes. Evidence on extra-time programmes is mixed. Studies of double-dose programmes in ninth grade (Year 10) English (Nomi, 2015) found positive impacts on multiple outcomes including course grades, standardised tests, and graduation rates, with larger impacts for initially lower-performing students. One study in middle school (Dougherty, 2013) found that while there were immediate benefits of double-dose reading in sixth grade (Year 7), the impacts had disappeared by eighth grade (Year 9). However, reviews of studies of extra-time programmes by Kidron & Lindsay (2014) and Zeif, Lauver, & Maynard (2006) failed to find positive effects.

2.   **Programmes incorporating technology.** A number of widely used secondary reading programmes incorporate technology. These programmes vary greatly. Most provide self-paced instruction at students' reading levels, with immediate feedback and rewards as students progress. Some of these, such as READ 180 and Passport Journeys, integrate whole-group and small-group instruction with computer-based practice and instruction. Other programmes use technology as a means to deliver instruction and practice that is additional to the normal classroom curriculum, such as the eBooks and activities included with Accelerated Reader and iLit.

19

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; Borman et al., 2008). Programmes using technology were distributed among categories according to what they do, rather than in a separate technology category, but we also computed mean effect sizes for all technology approaches.

Results

A total of 69 studies evaluating 51 different programmes met the criteria of this review (note: when two distinct programmes were compared to control groups and reported in the same article, they counted as two "studies"). As a group, the studies were of very high methodological quality. 62 (90%) used random assignment, and only 7 (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). Table 1 summarises effect sizes and other information for the ten categories and the six cross-cutting factors.

It is important to note that because of the substantial representation of large, cluster randomised trials, effect sizes for all programmes are lower than most readers expect to see. This is characteristic of such research designs, as both random assignment and large sample size contribute to small effect sizes. For example, Cheung & Slavin (2016) computed average effect sizes according to various methodological factors across 645 studies of reading, math, and science accepted by the Best Evidence Encyclopedia (www.bestevidence.org). For randomised evaluations with sample sizes over 250, the mean effect size was +0.12. Therefore, effect sizes above this level might be considered above average for this design.

The text briefly describes individual studies. Tables 2 to 11 group studies according to their main components.

Tutoring Interventions

One-to-one or small-group tutoring is a widely used and effective intervention for struggling readers in elementary schools (Slavin, Lake, Davis, & Madden, 2011), but is rarely used in secondary schools. All seven of the qualifying tutoring studies (Table 2) were done in England, as a result of the EEF funding initiative focused on helping struggling students enter secondary school with adequate reading skills.

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 out loud, discussing the text, and linked writing. A study of Catch Up® Literacy in Year 6 (Rutt, Kettlewell, & Bernardenelli, 2015) found a marginally significant difference favouring the tutored students (ES=+0.16, p=.08).

Perry Beeches provides struggling readers with one hour of one-to-one tutoring every two weeks. Coaches tailor activities according to students' needs. A study of Perry Beeches by Lord, Bradshaw, Stevens, & Styles (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 teaching assistants. Sibieta (2016) evaluated two very similar variations, and found a mean effect size of +0.42.

Small-group tutoring.

Butterfly Phonics. Butterfly Phonics uses formal phonics instruction, understanding the global aspects of a text, and class discussion of text meaning to improve reading comprehension.

The programme is delivered to groups of 6-8 students by a trained practitioner and an assistant. Merrell & Kasim (2015) evaluated Butterfly Phonics, and found an effect size of +0.30 (p<.001).

Rapid Phonics combined with Sound Discovery. Rapid Phonics and Sound Discovery is a small-group tutoring programme 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 while control students continued their schooling as usual. In the evaluation (King & Kassim, 2015), the effect size was not significant (ES= -0.05, n.s.).

Taken together, the overall weighted effect size for tutoring programmes provided by paid adults was +0.24 (p<.001). Effect sizes were +0.28 (p<.001) for three one-to-one programmes and +0.14 (n.s.) for two small-group programmes.

Volunteer tutoring.

TextNow utilises volunteer coaches to provide daily 20-minute sessions focusing on encouraging students to read for pleasure. A randomised trial of TextNow in England (Maxwell et al., 2014) found no 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 significant effects (ES = -0.02, n.s.) in a randomised experiment (Lloyd et al., 2015).

Cooperative Learning Approaches.

Cooperative learning methods, in which students work in small groups to help each other grow in reading skills, are widely used in reading and in many other subjects in the elementary grades. Six studies of two cooperative learning programmes are shown in Table 3.

The Reading Edge, adapted from a programme called Student Team Reading, is a cooperative learning programme 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 in the US of The Reading Edge, the weighted mean effect size was +0.15 (Slavin, Chamberlain, Daniels, & Madden, 2009; Stevens & Durkin, 1992a, b). The outcomes in the Slavin, Chamberlain, Daniels, & Madden (2009) study were significantly positive (ES=+0.15, p<.05).

Collaborative Strategic Reading (CSR*)* teaches reading comprehension strategies to students working in small cooperative learning groups. The weighted mean effect size across three CSR studies in the US was +0.05 (Denver Public Schools, 2016; Vaughn et al., 2011, 2013). Adding the findings of the CSR studies to those of the three Reading Edge studies, the weighted mean effect size for all cooperative learning studies was +0.10 (p<.05).

Whole-School Approaches

Whole-school approaches provide professional development to teachers across entire schools or grade levels. Two of these approaches emphasise generic teaching methods, school organisation (usually including teacher teams), and schoolwide approaches designed to make entire schools more focused on effective strategies for improving achievement and social-

emotional development. Five others provided professional development to all teachers of academic subjects, but did not emphasise organisational elements. All seven are shown in Table 4.

Whole-school approaches with organisational elements.

BARR (Building Assets Reducing Risks) is a whole-school reform approach focused on developmental, academic, and structural challenges during ninth grade (Year 10). BARR is used in all major subjects to attempt to increase student achievement by improving students' social-emotional skills, building positive student–teacher relationships, and solving non-academic barriers to learning, such as truancy and behaviour 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, maths, and science or humanities 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 professional development and coaching are provided to teachers and school leaders. In the US, Corsello & Sharma (2015) found a positive effect of BARR on reading (ES=+0.14, p<.01), and a larger study by Borman et al. (2017) also found significant positive effects (ES = +0.08, p<.05). The weighted mean across the two studies was +0.09.

Talent Development High School (TDHS) is a whole-school reform model for ninth graders (Year 10). Within TDHS, Strategic Reading and Student Team Literature (SR/STL) is the reading component. Students receive a "double dose" of reading and maths, amounting to 90 minutes a day for each subject. In SR/STL, students work in small, interdependent cooperative learning groups using structured partner discussion guides that provide background and vocabulary, and comprehension questions provide mini-lessons on specific comprehension strategies.

24

Balfanz, Legters, & Jordan (2004) carried out an evaluation of TDHS in high-poverty high schools in Baltimore, comparing to control schools that also provided double-dose reading. There was a non-significant effect size of +0.32 (significance was not attained because there were only six schools).

The weighted mean effect size for all three studies evaluating whole-school approaches with organisational elements was +0.09 (p<.05).

Whole-school approaches without organisational elements.

Teacher Effectiveness Enhancement Programme (TEEP) is a professional development approach for secondary teachers involving all major subjects. It includes formative assessment, thinking skills, cooperative learning, and effective use of technology. An evaluation in England 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 centres, handheld computers) to susspport their learning. All teachers were encouraged to assign partner reading in the humanities, science, and maths. A multi-year evaluation in the US found no significant effects (Simon et al., 2011; ES= -0.01, n.s.).

Project CRISS is a professional development 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. A US evaluation of CRISS (Kushman, Hanita, & Raphael, 2011) found no significant effects (ES=+0.05, n.s.).

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

25

Texas Technology Immersion Pilot (eTxTip) was a three-year evaluation of a technology immersion intervention covering English, maths, science, and humanities in grades 6-8 (Years 7-9). Contractors provided schools with a) wireless, mobile computing devices for every student and teacher, b) productivity, communication, and presentation software, c) online resources supporting state standards, d) online assessments linked to state standards, e) extensive professional development, and f) initial and ongoing technical support. A study by Shapley, Sheehan, Maloney, & Caranikas-Walker (2009) found a mean effect size of +0.06 (n.s.).

The weighted mean effect size for whole school strategies without organisational elements was 0.00. Across all seven studies of whole-school programmes, the mean was +0.06 (p<.05).

Writing-focused Approaches

Two programmes, summarised in Table 5, focus on teaching writing as a means of improving reading outcomes.

Pathway is a US professional development programme used primarily with Latino English learners in regular classes. Teachers learn how to teach cognitive strategies and process writing. A study by Olson et al (2012) found 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 (ERWC) is a programme for 12th graders designed to prepare them to pass the California Early Placement Test (EPT), used in the California State University system. The emphasis of the programme 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 significant positive effects of ERWC (ES= +0.13, p<.05).

26

The weighted mean across the three studies of writing-focused programmes was +0.13 (p<.001).

Content-focused Approaches

Seven studies in the US evaluated content-focused approaches that teach comprehension strategies for humanities, science, and other expository text (see Table 6).

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

Reading Apprenticeship Improving Secondary Education (RAISE) is a whole-class version of Reading Apprenticeship. A study of RAISE by Fancsali et al. (2015) found non-significant positive effects (ES=+0.14, n.s.).

iRAISE is a form of Reading Apprenticeship that provides professional development online. Jaciw, Schellinger, Lin, Zacamy, & Toby (2016) found no positive effects of this approach (ES = 0.00, n.s.).

Reading Apprenticeship Academic Literacy (RAAL) is an adaptation of Reading Apprenticeship designed for struggling readers, taught during an extra daily reading period. A study of RAAL by Somers et al. (2010) found 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.

Every Classroom, Every Day (ECED) is a structured literacy curriculum based on authentic expository texts. Early et al. (2015) found no significant effects of ECED on reading (ES=+0.06, n.s.).

ITSS (Intelligent Tutoring System for the Structure Strategy) is a web-based approach in which students are taught to comprehend nonfiction text by categorising text structures using key elements in the text to find the main idea, activating prior knowledge, supporting cognitive monitoring, and using graphic organisers and flow charts to summarise texts. ITSS is used during regular English classes with the assistance of a teaching assistant. Animated "tutors" model and guide learners. Students practise, take regular assessments, and proceed at their own pace through self-instructional units. A study by Wijekumar, Meyer, & Lei (2017) found a significant positive effect for 7th graders (Year 8) (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 analyse the text and answer critical questions. In a study by Simmons et al. (2014), the effect size was 0.01 (n.s.).

Across all seven studies of content-focused approaches, the weighted mean effect size was +0.08 (p<.05).

Vocabulary-Focused Approaches

Three programmes emphasised the teaching of vocabulary as a means of enhancing comprehension. Studies of these programmes are shown in Table 7.

Academic Language Instruction for All Students (ALIAS) is a vocabulary intervention designed to be used in regular English classrooms that including many EAL 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 ALIAS studies in the US was +0.07 (n.s.) (Lesaux, Kieffer, Faller, & Kelly, 2010; Lesaux, Kieffer, Kelley, & Harris, 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, Pare-Blagoev, & Snow (2016) evaluated reading effects of Word Generation in the US and found no significant differences on reading (ES= +0.05, n.s.).

Vocabulary Enrichment Intervention Programmeme (VEIP) combines vocabulary teaching, phonics, and connected text. Teachers are given the flexibility to use the components as needed by their students. An evaluation of VEIP by Styles et al. (2014) in England found no significant outcomes (ES = +0.06, n.s.).

Four studies of vocabulary-focused approaches had a weighted mean effect size of +0.06 (n.s.).

Strategy-focused Instruction

Strategy-focused instruction approaches are all variations of the Strategic Instruction Model (SIM; Deshler & Schumaker, 2005), a family of programmes all of which emphasise teaching students reading comprehension, decoding, and other reading objectives with step-by-step strategies. SIM studies, all in the US, are summarised in Table 8.

SIM: Xtreme Reading is the main version of the Strategic Instruction Model 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, Zaler, Kite, & Hussar, 2012), the mean effect size for Xtreme Reading was +0.09 (p < .01).

SIM: Content Literacy Curriculum (CLC). A large two-year study by Corrin et al. (2012) found non-significant positive effects on reading (ES=+0.09, n.s.).

SIM: Learning Strategies Curriculum (LSC). A study by Cantrell, Almasi, Rintamaa, & Carter (2016) found significant positive effects of LSC on reading (ES= +0.10, p<.05).

SIM: Fusion Reading A study by Schiller at al. (2012) found non-significant positive effects of Fusion Reading (ES= +0.07, n.s.).

SIM: Adolescent Literacy Model (ALM) was evaluated by Cantrell, Almasi, Carter, & Rintamaa (2011) in Kentucky middle and high schools. They reported a non-significant effect size of +0.10.

The weighted mean effect size across 8 studies of five SIM variations was +0.09 (p<.001).

Personalisation Approaches

Personalisation approaches provide content appropriate to students' reading levels, usually using computers. In some programmes, there is a strong element of choice, where students may select texts of interest to them. Six studies, summarised in Table 9, fell into this category.

The 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 (such as summarising, questioning, predicting, or visualising) and

prompt students to use them. In a study in the US by Drummond et al. (2011), the average effect size was +0.01 (n.s.).

Schoolwide Enrichment Model – Reading (SEM-R) exposes students to a variety of books. They spend time independently reading 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 in the US by Little, McCoach, & Reis (2014) found non-significant positive effects (ES=+0.10, n.s.).

Achieve3000 is an online literacy programme that provides non-fiction reading content and teaches metacognitive skills to improve the comprehension of informational texts. In an evaluation in the US by Shannon & Grant (2015), the effect size was +0.29 across grades 6 and 9 (Years 7 and 10) (p<.05).

SuccessMaker is an adaptive K-8 (Years 1-9) computer-based reading programme. It provides individualised reading activities, game-like environments, interactive aids, and a reporting system to inform teachers on student progress. A study in the US by Gatti (2011) found a nonsignificant effect size for seventh graders (Year 8) of +0.11 (n.s.).

Accelerated Reader is a widely used U.S. programme, but the only qualifying evaluation in secondary reading took place in England (Gorard, Siddiqui, & See, 2015). It provides students with a wide range of books at their reading level, determined by an on-line test. On-line comprehension tests are provided for each book, and students can earn points based on completing many books at a high readability level. The Gorard et al. (2015) evaluation found a significant effect size of +0.24 (p<.05).

iLit is a digital instruction approach for struggling readers. Students choose among more than 500 ebooks and work on vocabulary and comprehension strategies. Students work

independently, keeping on-line journals, answering questions, and discussing books in groups. A two-year study in the US of iLit by Gatti (2016) found an effect size of +0.09 (n.s.).

Across all studies of personalisation approaches, the weighted mean effect size was +0.13 (p<.05).

Group/Personalisation Rotation Approaches

Seventeen studies of nine programmes, mostly using technology, rotate students through group instruction from the teacher, personalised work at students' own reading levels, and (often) cooperative learning. They are summarised in Table 10.

READ 180 is an instructional model used 90 minutes each day with struggling readers. It combines 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 lessons, 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 in the US (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. It provides student texts and supportive on-line materials focused on building reading, writing, and comprehension skills, including whole-class and small-group discussions, teaching of metacognitive skills such as graphic organisers, and collaborative projects. Sivin-Kachala & Bialo (2012) found non-significant positive effects of Expert 21 (ES= +0.15, n.s.) in the US.

System 44 is a version of READ 180 for adolescent readers who have not mastered basic phonics and decoding skills. The programme focuses on decoding, fluency, and comprehension. In daily 60-minute lessons, the teacher gives 5-10 minutes of whole-class instruction, the students spend 25-30 minutes working in small groups or individually, and they then receive 20-25 minutes of computer-delivered instruction. Like READ 180, System 44 is always used during supplemental reading time, but in a study in the US by Beam, Faddis, & Hahn (2011), the control group also had additional reading time, so there were no time differences. Outcomes were significantly negative on the Test of Silent Reading Efficiency and Comprehension (TOSREC) (ES= -0.24, p<.05) and non-significantly negative on CST (ES=-0.04), for a mean of ES=-0.14. A study in the US that did provide supplemental reading time, by Beam & Faddis (2012), found a significant positive impact (ES=+0.20, p<.05). Across the two studies, the mean ES was +0.03.

Passport Reading Journeys is a supplemental literacy curriculum that provides fifteen two-week sequences of lessons mixing whole-class and smallgroup lessons as well as individualised 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 in the US (Dimitrov et al, 2012; Schenck et al, 2012; Vaden-Kiernan et al., 2012) the weighted mean effect size was +0.07. The Vaden-Kiernan et al. (2012) study found significant positive effects on the GRADE (ES=+0.27, p<.05), but not the LEAP (ES=-0.01), for a mean of +0.12.

Comprehension Circuit Training (CCT) uses content delivered on tablet computers to teach reading comprehension skills. Following video instruction, students work with a partner to practise lesson content. Students cycle through four major components, focusing on vocabulary skills, pre-reading, reading of ebooks, and comprehension quizzes. Across two studies in the US

(Fogarty et al., 2014, 2016), the mean effect size for CCT was +0.13. In the Fogarty et al. (2016) study, significant positive effects were found on the TOSREC (ES=+0.24, p<.05), but not other measures.

Prentice Hall Literature combines printed textbooks with online components. Online material includes vocabulary games, audios, and videos. A study in the US by Eddy, Ruitman, Hankel, & Sloper (2010) found non-significantly negative impacts of Prentice Hall Literature (ES= -0.10).

Strategic Adolescent Reading Intervention (STARI) uses core novels and other engaging texts within thematic units, fluency passages, comprehension instruction, reciprocal teaching strategies, as well as student discussion and debate. Kim et al. (2017) conducted a randomised trial of STARI with low-achieving middle school students in the Northeastern United States and while the overall effects of the programme were non-significantly positive, (ES = +0.15, p = n.s.), the programme did demonstrate significantly 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 in the US by Deussen et al. (2012) found non-significant positive effects on reading (ES= +0.10, n.s.).

Reading Intervention through Strategy Enhancement (RISE) relies on teachers' capacity to build an effective curriculum for struggling readers. During RISE classes, students are given the opportunity to read independently, to work in small groups, and to receive whole group lessons. A study in the US of RISE by Lang et al. (2009) found positive effects on reading for

"moderate risk" students (ES= +0.27, p<.04), but not for "high risk" students (ES= -0.06, n.s.). The weighted mean effect size was +0.16 (p<.05).

Across all 17 studies of rotation models, the weighted mean effect size was +0.09 (p<.001).

Intensive Approaches

Intensive approaches are designed to help students catch up on decoding and word study skills, as well as basic comprehension strategies, which other students would have mastered in primary school. These studies are summarised in Table 11.

REWARDS is a one-year supplemental intervention given by trained teachers five times a week to develop reading skills with low achievers. An evaluation in the US of REWARDS by Newman, Kundert, Spaulding, White, & Gifford (2012) found a mean effect size of +0.09 (n.s.).

Kentucky Cognitive Literacy Model (KCLM) is an intervention for struggling readers focused on teaching comprehension strategies, vocabulary, study skills, and writing. A study of the model by Cantrell, Carter, & Rintamaa (2012) found no 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, on narrative structure and writing skills, and spelling. Lang et al. (2009) evaluated REACH in the US and found a mean effect size of -0.02 (n.s.).

Across the three studies of the intensive approaches, the weighted mean effect size was 0.00 (n.s.).

Differences by Cross-Cutting Factors

All studies included in random effects models were used to explore several cross-cutting features that differed within treatment categories.

35

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. The control group was typically participating in elective lessons, such as art or music, or they were in revision classes, 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 to 20).

An analysis that compared studies providing extra reading periods and studies that did not found no significant differences. Twenty-seven studies of programmes providing extra reading time had a mean weighted effect size of +0.09, while 42 studies of programmes providing no extra time had a weighted mean effect size of +0.10.

Use of technology. We compared programmes making extensive use of technology to those that made little or no use of technology. Programmes using technology (n=23) obtained non-significantly lower effect sizes (ES=+0.08) than those that did not use technology (n=46, ES=+0.10). A few individual programmes making extensive use of technology, such as Achieve 3000 (ES=+0.29), Accelerated Reader (ES=+0.24), and ITSS (ES=+0.18) did report significant positive impacts, but these were exceptions.

Early vs late secondary. We tested the difference in outcomes between programmes used in the early years of secondary (Years 7-9) with those used later on (Years 10-13). Weighted mean effect sizes were nearly identical in the early (n=44 studies, ES=+0.10) and later years (n=30 studies, ES=+0.10).

Struggling readers and English learners. Some qualifying studies served only struggling readers, usually those performing at least two years below their Year level. Others served all students, though in most cases the schools involved had high levels of poverty. We compared

outcomes for studies focused only on struggling readers (n=36) to those focused on all students (n=36). Effect sizes were nearly identical, +0.10 for struggling readers and +0.10 for all students.

None of the qualifying studies reported outcomes separately for pupils with English as an Additional Language (EALs), but two programmes focused on EALs and showed promising outcomes: Pathway (mean ES=+0.08), and ALIAS (mean ES=+0.06). Among the two studies of Pathway, one found statistically significant positive effects.

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 (n=7 studies, ES=+0.13). This difference was not significant, though similar to differences reported by Cheung & Slavin (2016). Effect sizes were non-significantly 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).

Discussion

This review of rigorous research on programmes designed to enhance the reading of students in secondary schools found that most studies meeting inclusion criteria had relatively small effects on student reading. Weighted mean effect sizes for all categories are shown in Table 1. The Table also shows mean differences in effect sizes for cross-cutting factors (e.g., technology vs. no technology, early and late secondary school). As noted earlier, effect sizes for large, cluster randomised experiments are typically much lower than those for smaller or quasi-experimental studies, though their outcomes are more reliable and more likely to replicate than other designs (Lipsey & Wilson, 2001). Cheung & Slavin (2016) found a mean effect size of +0.12 for such

37

studies, so large randomised experiments with larger effect sizes than this may be considered above average for their category.

A few quite different categories contained programmes with more positive impacts. One of these was tutoring programmes, all done in England, in which paid adult tutors worked with groups of one to four students (or in one case, two adults to 6-8 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 primary reading (Slavin, Lake, Davis, & Madden, 2011; Wanzek et al., 2013). Tutors are able to thoroughly personalise instruction to students' individual needs, and to build personal relationships with them. However, 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 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 behaviour. By having students work in teams, with team recognition based on the achievement gains of all team members, teammates encourage each other's efforts, explain ideas to each other, and have opportunities to ask others for help (Slavin, 1995; Roseth et al., 2008).

However, three studies of another cooperative learning approach, Collaborative Strategic Reading, found small impacts (weighted mean ES= +0.05). Combining across all six studies of cooperative learning, the mean effect size was +0.10.

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

38

Three studies of whole-school designs with organisational elements such as teacher teams showed particular promise. BARR (Building Assets, Reducing Risk), a whole-school model, has been evaluated in two high-quality randomised experiments. BARR organises ninth grade (Year 10) 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 programme emphasises social-emotional development and relationships among teachers and students. A study of Talent Development High School, which is also a whole-school approach with teacher teaming, had an impressive but non-significant effect size of +0.32. However, other whole-school approaches lacking organisational elements did not show positive outcomes.

Two programmes with a strong emphasis on writing found positive effects on reading. The Expository Reading and Writing Course (ERWC), a programme designed to help twelfth graders (Year 13) prepare for the test they will take in their first year if they attend California State Universities, reported a significant positive effect on this test, with an effect size of +0.13. Pathway, a professional development programme primarily for teaching students with English as an Additional Language, also showed significant effects across two studies (ES = +0.08). The mean for this category was +0.13.

One family of approaches with mostly positive effects was the Strategic Instruction Model (SIM), which teaches struggling adolescent readers step-by-step strategies for phonics, comprehending, writing, note-taking, and other skills. Eight studies evaluated SIM variations and three of these found significant positive effects. The outcomes of all eight studies averaged +0.09.

Another interesting family of programmes was Reading Apprenticeship (Greenleaf et al., 2011), which focuses on comprehension and writing strategies. The only Reading Apprenticeship

variation to have significant positive effects was Reading Apprenticeship Academic Literacy (RAAL; Somers et al., 2010), and the mean across the four studies was +0.07.

One of the most surprising findings of this review is the lack of additional positive effects of providing an additional class period for reading each day, compared to providing extra support in class time. Programmes focused on improving teachers' practices during regular class periods produced virtually identical impacts as programmes that also focused on improving teachers' practices but added a daily period of instruction. The finding matches the conclusions of reviews by Kidron & Lindsay (2014) and Zief, Lauver, & Maynard (2006).

The failure to find a larger impact of additional instructional time in reading, compared to providing extra support in class time, 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 PE, in most cases) and were not motivated to once again work on material they had difficulty with in primary school.

Similarly, it was surprising to find that there was no impact of programmes emphasising technology, although there were a few individual exceptions. An earlier review by Cheung & Slavin (2013) did find mostly positive, though small, outcomes of technology for reading outcomes. It may be that teachers are still not comfortable with technology. One indication of this might be that of the few technology-focused programmes that did find 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. Still, given the enormous emphasis and expenditure on technology common today, this finding is disturbing.

There are a few commonalities among programmes that achieved positive outcomes. There are several of these worth mentioning, though they are far from conclusive.

One interesting commonality was that programmes with positive outcomes tended to emphasise student motivation, student-to-student and student-to-teacher relationships, and social-emotional learning. Positive examples include cooperative learning (at least The Reading Edge), BARR and Talent Development, which focus on relationships and social-emotional learning, and tutoring, which provides immediate feedback and potentially close teacher–student relationships.

Another factor seen in many successful programmes is personalisation. This is most obvious in tutoring and in some technology approaches. Cooperative learning also provides a form of personalisation, as students help each other succeed.

The positive impacts of both writing-focused approaches, ERWC and Pathway, replicates earlier research summarised by Graham & 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 programmes can appear in many categories. As research in secondary reading continues, it will be useful and instructive to vary programme components to learn which generic approaches most enhance student outcomess.

The research reviewed here identifies specific proven programmes and outlines 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 secondary school students.

References*

*Balfanz, R., Jordan, W., & Legters, N. (2004). *Catching Up: Impact of the Talent Development ninth grade instructional interventions in reading and mathematics in high-poverty high schools*. Center for Research on the Education of Students Placed at Risk, Johns Hopkins University.

Baye, A., Lake, C., Inns, A., & Slavin, R. E. (2019). Effective reading programs for secondary students. *Reading Research Quarterly, 54*(2), 133-166.

*Beam, M. & Faddis, B. (2012). *Evaluation of System 44. Final report*. Portland, OR: RMC Research Corporation.

*Beam, M. Faddis, B. & Hahn, K. (2011). *Evaluation of System 44. Final report*. Portland, OR: RMC Research Corporation.

Biancarosa, G. (2012). Adolescent literacy: More than remediation. *Educational Leadership*, *69*(6), 22–27.

Biancarosa, C., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy.* 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, T., Bos, J., O'Brien, B., Park, S. J., 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 - results from a randomized controlled trial.* Washington, DC: SREE.

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.

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.* Washington, DC: USDOE.

*Cantrell, S. C., Almasi, J. F., Carter, J. C. & Rintamaa, M. (2011). *Striving Readers final evaluation report: Danville, KY.* Lexington, KY: 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. doi:10.1080/00220671.2014.917258.

*Cantrell, S. C., Carter, J. C., & Rintamaa, M. (2012). *Striving Readers Cohort II Evaluation Report: Kentucky*. Lexington, KY: 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.

Cheung, A., & Slavin, R. (2016). How methodological features affect effect sizes in education. Educational Researcher, 45 (5), 283-292. DOI: https://doi.org/10.3102/0013189X16656615.

Cockroft, C., & Atkinson, C. (2017). The effectiveness of affective literacy interventions for adolescents: a review of the literature. *Educational Psychology in Practice: theory, research and practice in educational psychology*, *3*(1), 29-49.

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

*Corrin, W., Lindsay, J. J., Somers, M-A., Myers, N. E., Myers, C. V., Condon, C. A., & Smith, J. K. (2012). *Evaluation of the Content Literacy Continuum: Report on program impacts,*

program fidelity, and contrast. (NCEE2013-4001). Washington, DC: National Center for

Education Evaluation and Regional Assistance, IES, 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*. Napa, CA:

Corsello Consulting.

de Boer, H., Donker, A. S., & van der Werf, M. P. (2014). Effects of the attributes of educational

interventions on students' academic performance: A meta-analysis. *Review of Educational

Research*, *84*(4), 509–545. doi: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 Press.

*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.

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. doi 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

Corporation.

44

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 Society for Research on Educational Effectiveness, Washington, D.C.

*Drummond, K., Chinen, M., Duncan, T., Miller, H., Fryer, L., Zmach, C., & ... National Center for Education Evaluation and Regional Assistance, (2011). *Impact of the Thinking Reader[R] 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, IES, U.S. Department of Education.

*Early, D., Berg, J. K., Alicea, S., Si, Y., Aber, J. L., Ryan, R. M., & Deci, E. L. (2015). 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. doi: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. doi:10.3102/0034654308325998

*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 Corporation.

*Fancsali, C., Abe, Y., Pyatigorsky, M., Ortiz, L., Hunt, A., Chan, V., … Jaciw, A. P. (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 Inc.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35*(1), 39–50. doi: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 and Practice, 27*(1), 21-32. doi:10.1111/j.1540-5826.2011.00347.x

*Fogarty, M., Clemens, N., Simmons, D., Anderson, L., Davis, J., Smith, A., ... & Oslund, E. (2016). Impact of a technology-mediated reading intervention on adolescents' reading comprehension. *Journal of Research on Educational Effectiveness, 10*, 326-353. Retrieved from

http://www.tandfonline.com/doi/abs/10.1080/19345747.2016.1227412?journalCode=uree 20.

*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, 2*6(3), 425-449. doi:10.1007/s10648-014-9270-6

*Fong, A., Finkelstein, N., Jaeger, L., Diaz, R., & Broek, M. (2015). *Evaluation of the expository Reading and Writing Course: Findings from the Investing in Innovation development grant.* San Francisco, CA: WestEd.

*Gatti, G. (2011) *Pearson SuccessMaker Reading efficacy study 2010-2011 final report.* Pittsburgh, PA: Gatti Evaluation, Inc.

*Gatti, G.G. (2016). *iLit 2013-2015: 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. doi:10.1353/foc.2012.0011

* Gorard, S., Siddiqui, N., & See, B. H. (2015). *Accelerated reader: Evaluation report and executive summary*. London, England: EEF.

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. doi: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, 0*(0), 1–26. doi: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. doi:10.3102/0002831210384839

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.* Washington, DC: U.S. Department of Education.

*IEE (2016). *Teacher Efectiveness Enhancement Programme: Evaluation report and executive summary.* York, England: Author.

*Jaciw, A. P., Schellinger, A. M., Lin, L., Zacamy, J., & Toby, M. (2016). *Effectiveness of internet-based reading apprenticeship Improving science education (iRAISE).* 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. doi: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* (No. NCEE #2008-4027). Washington, D.C.: National Center for Education Evaluation and Regional Assistance, IES, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc

Kidron, Y., & Lindsay, J. (2014). *The effects of increased learning time on student academic and nonacademic outcomes: Findings from a meta-analytic review* (REL 2014-2015). Washington, DC: U. S. Department of Education, IES.

*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. doi:10.1002/rrq.171

*King, B., & Kasim, A. (2015). *Rapid Phonics: Evaluation report and executive summary*. Millbank, England: 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*

48

NCEE 2011-4007. Washington, DC: National Center for Education Evaluation and Regional Assistance.

*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. doi: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. doi:10.3102/00028312038004837

*Lawrence, J., Francis, D., Pare-Blagoev, J., & Snow, C. (2017). The poor get richer: Heterogeneity in the efficacy of a school-level intervention for academic language. *Journal of Research on Educational Effectiveness, 10,* 767-793. doi:10.1080/19345747.2016.1237596

Lee, C. D., & Spratley, 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. doi: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. doi: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. doi:10.1177/1932202X14549250

*Lloyd, C., Edovald, T., Kiss, Z., Morris, S., Skipp, A., & Ahmed, H. (2015). *Paired Reading: Evaluation report and executive summary.* London, England: Education Endowment Foundation. Retrieved from http://socialwelfare.bl.uk/subject-areas/services-activity/education-skills/natcen/175224Paired_Reading.pdf

*Lord, P., Bradshaw, S., Stevens, E. & Styles, B. (2015). *Perry Beeches Coaching Programmeme: Evaluation Report and Executive Summary*. London, England: 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, England: Education Endowment Foundation. Retrieved from http://www.paulconnolly.net/publications/EEF_TextNow_2014.pdf

*Meisch, A., Hamilton, J., Chen, E, Quintanilla, P., Fong, P., Gray-Adams, K., … Thornton, N. (2011) *Striving Readers study: Targeted & whole-school interventions – Year 5*. Rockville, MD: Westat.

*Merrell, C., & Kasim, A. (2015). *Butterfly Phonics: Evaluation report and executive summary*. London, England: 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. doi:10.3102/0162373716638446

National Center for Education Statistics (NCES) (2016). *National Assessment of Educational*

Progress. Washington, DC: Author.

Nelson, N., & Calfee, R. (1998). The reading-writing connection. In N. Nelson & R. Calfee (Eds.), *Ninety-seventh yearbook of the National Society for the Study of Education* (Part II, pp. 1-52). Chicago, IL: National Society for the Study of Education.

*Newman, D. L., Kundert, D. K., Spaulding, D. T., White, S. P., & Gifford, T. A., (2012). *Striving Readers Local Evaluation*. Albany, NY: University at Albany, State University of New York at Albany.

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. doi.10.1016/j.ssresearch.2014.12.012

OECD (2016). *PISA 2015 results (Volume I): Excellence and equity in education.* Paris, France: Author. doi 10.1787/9789264266490-en

OECD (2013). *OECD skills outlook 2013: First results from the survey of adult skills*. Paris, France: Author. doi: 10.1787/9789264204256-en

*Olson, C., Kim, J., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D., ... Land, R. (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. doi: 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*, 7-12.. doi:10.1037/edu0000095.

51

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*, 355-379.. doi 10.1007/s11145-017-9788-3

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. doi: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,* pp. 333-356). Hoboken, NJ: Wiley.

R Core Team (2016). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*, 2nd ed. Thousand Oaks, CA: Sage Publications, Inc.

Roseth, C., Johnson, D., & Johnson, R. (2008). Promoting early adolescents' achievement and peer relationships: The effects of cooperative, competitive, and individualistic goal structures. *Psychological Bulletin, 134*(2), 223–246. doi:10.1037/0033-2909.134.2.223

*Rutt, S., Kettlewell, K., & Bernardenelli, D. (2015). *Catch Up® Literacy: Evaluation report and executive summary.* Millbank, England: Educational Endowment Foundation.

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: RMC Research Corporation, Center on Instruction.

*Schenck, A., Feighan, K., Coffey, D., & Rui, N. (2011). *Memphis Striving Readers project: Evaluation report, Year 4.* New York, NY: Research for Better Schools & RMC Research Corporation.

*Schenck, A., Jurich, S., Frye, M., Lammert, J., Sayko, S, Najera, K. …Willard, T. (2012) *Evaluation report/Impact study: Virginia Striving Readers Intervention Initiative.* Arlington, VA: RMC Research Corporation.

*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.

*Shannon, L., & Grant, B-J. (2015). *An evaluation of the Achieve 3000 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, TX: Texas Center for Educational Research.

*Sibieta, L. (2016). *REACH: Evaluation report and executive summary*. London, England: Education Endowment Foundation.

*Simmons, D., Fogarty, M., Oslund, E., 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*, 309-330. doi:10.1080/19345747.2013.836766

*Simon, A., Tunik, J., Alemany, J., Zhu, J., Zacharia, J., Ramsay, L., …Mendes, R. (2011). *Chicago Public Schools Striving Readers Initiative: Year four evaluation report.* New York, NY: Metis Associates.

*Sivin-Kachala, J., & Bialo, E. (2012). *Program evaluation research on Expert 21 in grades 6-8 for Scholastic Education.* New York: IESD.

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

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice (2nd Ed.).* Boston: Allyn & Bacon.

Slavin, R.E. (2015). Cooperative learning in schools. In J. Williams (Ed.), *International encyclopedia of the social and behavioral sciences (2nd ed.),* 881-886. Amsterdam: Elsevier.

* Slavin, R., Chamberlain, A., Daniels, C., & Madden, N. (2009). The Reading Edge: a randomized evaluation of a middle school cooperative reading program. *Effective Education, 1*(1), 13-26.

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

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

*Somers, M. A., Corrin, W., Sepanik, S., Salinger, T., Levin, J., Zmach, C., & 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.

*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: Brown University: The Education Alliance.

*Stevens, R. J., & Durkin, S. (1992a, b). *Using student team reading and student team writing in middle schools: Two evaluations.* Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education.

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.* London, England: EEF.

*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.

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). Teacher professional learning and development: Best evidence synthesis iteration (BES). Wellington, New Zealand: Ministry of Education.

Uccelli, P., Galloway, E. P., 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

*Vaden-Kiernan, M., Caverly, S., Bell, N., Sullivan, K., Fong, C., Atwood, E.... Jones, D. (2012). *Louisiana Striving Readers: Final evaluation report*. Austin, TX: SEDL.

Vaughn, S., & Fletcher, J. M. (2012). Response to intervention with secondary school students with reading difficulties. *Journal of Learning Disabilities*, *45*(3), 244–256.

*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. doi: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. doi: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.

Vygotsky, L. S. (1978). *Mind in society*. (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.

Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review, 36*, 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. doi: 10.3102/0034654313477212

What Works Clearinghouse (2017). *Procedures and standards handbook 4.0.* Retrieved from http://ies.ed.gov/ncee/wwc/Protocols.

What Works Clearinghouse. (2016). *Adolescent reading.* Washington, DC: U.S. Department of Education, Institute of Education Sciences.

*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*, 741-760. doi 10.1037/edu0000168

Zeif, G., Lauver, S., & Maynard, R. (2006). *The impacts of afterschool programs on student outcomes: A systematic review for the Campbell Collaboration.* Retrieved August 3, 2017 from https://www.campbellcollaboration.org/library/impacts-of-after-school-programmes-on-student-outcomes.html.

Table 1. Summary of effect sizes by category.

Program	k	Effect Sizes	Confidence Interval		Q	I^2	τ^2
			Low	High			
Table 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
Table 3: Cooperative Learning	6	+0.10*	0.00	+0.20	1.74	0.00	0.00
Table 4: Whole School	8	+0.06*	0.00	+0.13	3.22	0.00	0.00
With Organisational Elements	3	+0.09*	+0.02	+0.17	0.73	0.00	0.00
Without Organisational Elements	5	0.00	-0.12	+0.11	0.64	0.00	0.00
Table 5: Writing-Focused	3	+0.13***	+0.08	+0.17	0.64	0.00	0.00
Table 6: Content-Focused	7	+0.08*	+0.01	+0.15	3.32	0.00	0.00
Table 7: Vocabulary-Focused	4	+0.06	-0.06	+0.18	0.20	0.00	0.00
Table 8: Strategy-Focused Instruction	8	+0.09***	+0.04	+0.14	4.06	0.00	0.00
Table 9: Personalisation	6	+0.13*	+0.02	+0.23	3.49	0.00	0.00
Table 10: Group/Personalisation Rotation	17	+0.09***	+0.04	+0.13	9.85	0.00	0.00
Table 11: Intensive	3	0.00	-0.10	+0.10	1.49	0.00	0.00
Cross-Cutting Factors							
Additional Reading Period (27, 42)	69	-0.01	-0.05	+0.03	52.69	0.00	0.00
Technology Application (23, 46)	69	-0.01	-0.06	+0.04	52.52	0.00	0.00
Struggling Readers/All Students (36, 36)	72	+0.01	-0.03	+0.05	56.58	2.54	0.00
Early/late Secondary School (44, 30)	74	0.00	-0.04	+0.05	57.74	0.00	0.00
Randomised/Quasi (62, 7)	69	-0.04	-0.10	+0.01	50.08	0.00	0.00
Clustered/Student-Level (34, 35)	69	-0.05	-0.10	0.00	49.30	0.00	0.00

*p<.05. **p<.01. ***p<.001.

Table 2
Tutoring Interventions for Struggling Readers

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Post-test	Effect sizes	Overall effect size
One-to-One Tutoring								
Catch Up® Literacy								
Rutt et al. (2015)	SR	30 weeks	Students: 557 (286 E, 271 C)	Years 6-7	Students reading at least one year below year level from 15 schools mainly in urban areas across England. 21% FSM.	NGRT	+0.16	+0.16
The Perry Beeches Coaching Programme								
Lord et al. (2015)	SR	1 year	Students: 291 (149 E, 142 C)	Year 7	Students from 4 secondary schools in Birmingham who were reading at least one year below year level. 55% W, 20% ESL, 58% FSM.	GL Assessment - Progress in English	+0.36*	+0.36*
REACH (tutoring)								
Sibieta (2016)	SR	20 weeks	Students: 202 (70 REACH, 69 REACH + LC, 63 C)	Years 7-8	Lowest readers in 27 disadvantaged secondary schools in or near Leeds. 68% W, 32% non-white, 63% SEN, 24% EAL, 31% FSM.	NGRT REACH REACH + Language Comprehension	 +0.33* +0.51*	+0.42*
Small-Group Tutoring								
Butterfly Phonics								
Merrell & Kasim (2015)	SR	4 months	Students: 310 (161 E, 149 C)	Year 7	Students from 6 secondary schools in London who were reading at least one year below year level. 78% W, 16% AA, 35% SEN, 64% EAL, 51% FSM.	NGRT	+0.30*	+0.30*
Rapid Phonics combined with Sound Discovery								
King & Kasim (2015)	SR	12 weeks	Students: 178 (86 E, 92 C)	Years 6-7	Students from 22 primary and 13 secondary schools in Norfolk, who were reading at least one year below year level. 50% W, 50% SEN, 50% EAL, 50% FSM.	NGRT	-0.05	-0.05

Volunteer Tutoring

TextNow Transition Programme

Maxwell et al. (2014)	SR	15 weeks	Students: 391 (199 E, 192 C)	Years 6-7	Students reading below year level from 53 primary schools and 29 secondary schools across England. 16% EAL, 25% FSM.	NGRT	-0.06	-0.06

Cross-Age Tutoring

Paired Reading

Lloyd et al. (2015)	CR	16 weeks	Classes: 120 (58 E, 62 C) Students: 1306 (625 E, 681 C)	Year 7	10 schools in the North of England. 96% W, 20% FSM	NGRT	-0.02	-0.02

Notes for Tables 1-10

Design/Treatment: SR=Student Randomised, CR=Cluster Randomised, QE=Quasi Experiment, CQE=Cluster Quasi-Experiment, ARP=Additional Reading Period, TA=Technology Application

Measures: CAHSEE: California High School Exit Examination, CAT: California Achievement Test, CEM: Centre for Evaluation & Monitoring, CST-ELA: California Standards Test – English Language Arts, CTBS: Comprehensive Test of Basic Skills, DRP: Degree of Reading Power, EAL: English as a second language, ELA: English Language Arts, EPT: Early Placement Test (California), ETS: Educational Testing Service, FCAT: Florida Comprehensive Assessment Test, GORT: Gray Oral Reading Test, GSRT: Gray Silent Reading Test, GRADE: Group Reading Assessment and Diagnostic Evaluation, GMRT: Gates-MacGinitie Reading Tests, iLEAP, Louisiana State Reading Assessment, ISAT: Illinois Student Achievement Test, ISTEP+: Indiana State Test of Educational Proficiency, ITBS: Iowa Test of Basic Skills, KCCT: Kentucky Core Content Test, MAP : Measure of Academic Progress, MCAS: Massachusetts Comprehensive Assessment System, MEAP: Michigan Educational Assessment Program, MSP: Measurements of Student Progress state reading assessment, NGRT : New Group Reading Test (U.K.), NJASK: New Jersey State Test; NYS-ELA: New York State English Language Arts, NWEA: Northwest Evaluation Association, OAKS: Oregon Assessment of Knowledge and Skills, ORF: Oral Reading Fluency, RISE: Reading Inventory and Scholastic Evaluation, SAT 10: Stanford Achievement Test 10, SDRT-4: Stanford Diagnostic Reading Test 4, STAAR: State of Texas Assessments of Academic Readiness, SOL: Virginia Standards of Learning English/Reading, SWE: Sight Word Efficiency, TAKS: Texas Assessment of Knowledge and Skills, TCAP: Transitional Colorado Assessment Program, TOSREC, Test Of Silent Reading Efficiency and Comprehension, WJ III: Woodcock-Johnson III

Demographics: A=Asian, AA=African-American, H=Hispanic, W=White, FRL=Free/Reduced Lunch, ELL=English Language Learner, LD=Learning Disabilities, LEP=Limited English-proficient, SPED=Special Education. FSM=Entitlement to Free School Meals, EAL=English as an Additional Language, SEN=Special Educational Needs.

*p<.05 at the appropriate level of analysis (cluster or individual).

Table 3

Cooperative Learning Approaches

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Posttest	Effect sizes	Overall effect size
Cooperative Learning								
The Reading Edge/Student Team Reading								
Slavin, Chamberlain, Daniels, & Madden (2009)	SR	1 year	Students: 788 (405 E, 383 C) (2 cohorts)	6	2 Title I rural, mostly White middle schools in West Virginia and Florida. 90% W, 61% FRL, 15% SPED.	GMRT	+0.15[*]	+0.15*
Stevens & Durkin (1992a)	CQE	1 year	Schools: 5 (2 E, 3 C) Students: 3986 (1798 E, 2188 C)	6-8	High poverty, majority AA middle schools in Baltimore, Maryland.	CAT Comprehension	+0.34	+0.38
						CAT Vocabulary	+0.46	
Stevens & Durkin (1992b)	CQE	1 year	Schools: 6 (3 E, 3 C) Classes: 59 (20 E, 34 C) Students: 1223 (455 E, 768 C)	6	Middle schools in Baltimore, Maryland. 75% AA, 58% FRL.	CAT Comprehension	+0.13	+0.08
						CAT Vocabulary	-0.02	
Collaborative Strategic Reading (CSR)								
Denver Public Schools (2016)	CR	1 year	Schools : 16 Students : 5660 (3101 E, 2559 C) 3 cohorts	6-8	16 middle schools in Denver, Colorado. 62% H, 19% W, 11% AA, 30% ELL, 11% SPED, 76% FRL.	GMRT	+0.03	+0.03
						State Reading Test (TCAP)	+0.02	

						GMRT Comprehension	+0.12	
Vaughn et al. (2011)	CR	18 weeks	Classes: 61 (34 E, 27 C) Students: 782 (400 E, 382 C)	7, 8	6 middle schools from 3 school districts in Colorado and Texas. 43% W, 51% H, 52% FRL.	AIMSweb maze	-0.08	+0.04
						TOSREC	+0.07	
Vaughn et al. (2013)	CR	20 weeks	Classes: 48 (26 E, 22 C) Students: 472	7, 8	Same teachers and schools as in Vaughn et al. (2011). 51%W, 42%H, 6% LEP, 7% SPED.	GMRT	+0.10	+0.10
						TOSREC	+0.11	

Table 4
Whole-School Approaches

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Posttest	Effect sizes	Overall effect size
Whole-School with Organizational Elements								
BARR (Building Assets Reducing Risk)								
Borman et al. (2017)	SR	1 year	Students: 2172 (981 E, 1191 C)	9	6 schools in California (3), Maine (2), Minnesota (1). 71% minority, 21% ELL, 70% FRL.	NWEA	+0.08*	+0.08*
Corsello & Sharma (2015)	SR	1 year	Students: 495 (261 E, 234 C)	9	1 school in Southern California 52% W, 37% H, 11% AA, 17% ELL, 68% FRL.	NWEA	+0.14*	+0.14*
Talent Development High School (Strategic Reading and Student Team Writing)								
Balfanz et al., 2004	CQE	1 year	Schools: 6 (3 E, 3 C) Teachers: 20 E Students: 457 (257 E, 200 C)	9	High-poverty high schools in Baltimore, Maryland. 89% AA, 10% W, >90% FRL.	CTBS Terra Nova	+0.32	+0.32
Whole-School without Organizational Elements								
Teacher Effectiveness Enhancement Programme (TEEP)								
Institute for Effective Education (2016)	CR	1.5-2 years	Schools: 45 (23 E, 22 C) Students: 10,385 (5327 E, 8058 C) 12 cohorts	Year 9	Low-performing secondary schools across England. 30% FSM, 16% EAL.	CEM Insight-English	-0.04	-0.04
Chicago Striving Readers								
Simon et al. (2011)	CR/TA	1 to 3 years	Schools: 59 (29 E, 30 C) Students: 8127 (4074 E, 4053 C) (2 cohorts)	6-8	Middle schools across Illinois. 58% AA, 35% H, 9% SPED, 96% FRL.	ISAT Reading	-0.01	-0.01

Project CRISS

Kushman et al. (2011)	CR	1 year	Schools: 49 (23 E, 26 C) Students: 4959 (2460 E, 2499 C)	9	Schools in rural and urban fringe towns across 6 US Northwest states. 79% W, 15% FRL.	SDRT	+0.05	+0.05

eMINTS

						MAP		
Meyers et al. (2015)	CR/TA	3 years	Schools: 59 (20 E, 20 E+, 19 C) Students: 3295 (1208 E, 1216 E+, 871 C)	6-8	Rural middle schools across Missouri. 93%W.	eMINTS	-0.04	-0.06
						eMINTS + Intel	-0.08	

Texas Technology Immersion Pilot (eTxTip)

Shapley et al. (2009)	CQE/TA	2-3 years	Schools: 42 (21 E, 21 C) Students: 10,234 (4,767 E, 5,467 C) 3 cohorts	6-9	Rural, suburban, and urban middle schools across Texas. 70% H, 22% W, 7% AA, 15% LEP, 70% FRL.	TAKS	+0.06	+0.06

Table 5
Writing-focused Approaches

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Post-test	Effect sizes	Overall effect size
Pathway								
Kim et al. (2011); Olson et al. (2012)	CR	1 year	Teachers: 161 (79 E, 82 C) Students: 4459 (2200 E, 2259C) (2 cohorts)	6-11	15 schools (9 middle, 6 high) from a large school district in California. Eligible students: mainstreamed Latino ELLs. 95% H, 88% ELL, 79% FRL.	CST ELA	+0.07*	+0.07*
Olson et al. (2016)	CR	1 year	Teachers: 16 (9 E, 7 C) Students: 575 (313 E, 262 C)	10	Schools in Anaheim, California. 68% H, 18% A, 12% W, 20% ELL, 71% FRL.	CAHSEE	+0.19	+0.19
Expository Reading and Writing Course (ERWC)								
Fong et al. (2015)	QE	1 year	Students: 6618 (3309 E, 3309 C)	12	24 schools across California (15 urban, 3 rural, and 6 suburban). 45% H, 27% A, 24% W.	EPT	+0.13*	+0.13*

Table 6
Content-focused Approaches

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Posttest	Effect sizes	Overall effect size
Reading Apprenticeship								
Greenleaf et al. (2011)	CR	3 years	Schools: 78 (39 E, 39 C) Teachers: 111 (56 E, 55 C)	9-11	Biology teachers in 78 California schools. 48% H, 31% W, 19% ELL, 41% FRL.	DRP	-0.04	+0.03
						CST ELA	+0.10	
						Reading Comprehension	+0.13	
Reading Apprenticeship Improving Secondary Education (RAISE)								
Fancsali et al. (2015)	CQE	1 to 2 years	Schools: 42 (22 E, 20 C) Students: 10173 (5531 E, 4642 C) (2 cohorts)	9-12	High schools in California and Pennsylvania. 49% AA, 33% H, 10% ELL, 40% FRL.	ETS Literacy Achievement Assessment	+0.14	+0.14
iRAISE								
Jaciw et al. (2016)	CR	1 year	Schools: 26 Teachers: 69 (35 E, 34 C) Students: 1468 (751 E, 717 C)	9-12	High schools in Michigan and California. 73% W, 16% AA, 52% FRL.	ETS Literacy Assessment	0.00	0.00
Reading Apprenticeship Academic Literacy (RAAL)								
Somers et al. (2010)	SR/ARP	1 year	GRADE: Students: 2255 (1331 E, 924 C) (2 cohorts) State Test: Students: 1053 (2 cohorts)	9	Students from 17 high schools across multiple US districts who were reading 2-5 years below grade level. 31% H, 47% AA, 67% FRL.	GRADE Overall	+0.08	+0.10*
						Comprehension	+0.12*	
						Vocabulary	0.00	
						State Tests ELA	+0.15*	

Every Classroom, Every Day (ECED) /Literacy Matters							
Early et al. (2016)	CR/ARP	2 years	Schools : 20 (10 E, 10 C) Students: 8250 (3935 E, 4315 C)	9-10	20 high schools from 5 districts, 4 states (Arizona, Tennessee New York California). 51% H, 24% AA 22% ELL, 76% FRL.	State test ELA	+0.06	+0.06
ITSS (Intelligent Tutoring System for the Structure Strategy)							
Wijekumar, et al. (2017)	CR/TA	6-7 months	Classrooms: 108 (59 E, 49 C) Students: 2489	7	25 rural and suburban schools in the Northeast US. 92% W, 8% minority, 42% FRL.	GSRT	+0.18*	+0.18*
Content Knowledge-Building and Student-Regulated Comprehension Practices							
Simmons et al. (2014)	CR	1 semester	Classes: 65 (36 E, 29 C) Students: 786 (413 E, 373 C)	7-10	6 Title I schools (3 middle, 3 high) from 3 districts in one state in the Southwest US. 36% H, 31% AA, 71% FRL.	GMRT Comprehension	-0.01	-0.01

Table 7
Vocabulary-focused Approaches

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Posttest	Effect sizes	Overall effect size
Academic Language Instruction for All Students (ALIAS)								
Lesaux et al. (2010)	CR	18 weeks	Classes: 21 (13 E, 8 C) Students: 476 (296 E, 180 C)	6	7 middle schools in an urban Southwestern US district. 49% H, 73% ELL.	GMRT Comprehension	+0.15	+0.15
Lesaux et al. (2014)	CR	20 weeks	Teachers: 50 (25 E with their 37 classes, 25 C with their 39 classes) Students: 2082 (971 E, 1111 C)	6	14 urban middle schools in a large urban school district, California. 71% ELL, mainly Spanish speaking.	GMRT	+0.04	+0.04
						Comprehension	-0.04	
						Vocabulary		
Word Generation								
Lawrence et al. (2016)	CR	1 year	Schools: 44 (25 E, 19 C) Students: 8466 (4796 E, 3670 C)	6-8	US schools in 2 Northeast, 1 Western urban districts. 81% FRL.	GMRT		+0.05
						Reading Comprehension	+0.07	
						Vocabulary	0.00	
Vocabulary Enrichment Intervention Programme								
Styles et al. (2014)	SR	6 months	Schools: 11 Students: 570 (282 E, 288 C)	7	Students reading below year level from schools in England, 28% FSM.	NGRT	+0.06	+0.06

Table 8
Strategy-Focused Instruction

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Post-test	Effect sizes	Overall effect size
SIM: Xtreme Reading								
Somers et al. (2010)	SR/ARP	1 year	GRADE Students: 2329 (1341 E, 988 C) (2 cohorts) State Test Students: 1191 (2 cohorts)	9	Students from 17 high schools across multiple US districts who were reading 2-5 years below grade level. 31% H, 47% AA, 67% FRL.	GRADE Comprehension Vocabulary State Tests ELA	+0.05 +0.03 +0.08	+0.06
Faddis et al. (2011a) Middle school	SR	1 year	Students GRADE: 822 (401 E, 421 C) 4 Cohorts Students OAKS: 954 (472 E, 482 C)	7-8	6 Title 1 middle schools in Portland Oregon who were reading at least 2 years below grade level. 34% H, 23% AA, 27% SPED, 34% ELL.	GRADE Comprehension Vocabulary OAKS	+0.29* +0.32* +0.20* +0.12	+0.20*
Faddis et al. (2011b) High school	SR/ARP	1 year	GRADE: Students: 757 (355 E, 402 C) (4 cohorts) OAKS Students: 514 (260 E, 254 C) (4 cohorts)	9-10	Students from 4 Title I high schools in Portland, Oregon who were reading at least 2 years below grade level. 26% H, 35% AA, 24% SPED, 20% ELL.	GRADE Comprehension Vocabulary OAKS (10th grade only)	+0.12* +0.15* +0.07 +0.02	+0.08
Sprague et al. (2012)	SR/ARP	1 year	Students: 448 (223 E, 225 C) (5 cohorts)	9	Students from 5 Title I high schools in western Massachusetts who were reading between a 4th and 6th grade level. 75% minority, 22% SPED, 75% FRL.	SDRT-4	0.00	0.00

SIM: Content Literacy Curriculum (CLC)							
Corrin et al. (2012)	CR	1, 2 years	Schools: 28 (15 E, 13 C) Students: 1 year: 5011 (2975 E, 2036 C) 2 years: 4546 (2908 E, 1638 C)	9-10	High-poverty, low-achieving urban high schools across the US Midwest. 47% AA, 9% ELL, 11% SPED, 65% FRL.	**GRADE**	
						1 year	
						Comprehension	+0.06
						Vocabulary	+0.09	+0.09
						2 years	
						Comprehension	+0.10
						Vocabulary	+0.10
SIM: Learning Strategies Curriculum (LSC)							
Cantrell et al. (2016) Add'l period	SR/ARP	1 year	Grade 6 Students: 1135 (605 E, 530 C) (4 cohorts) Grade 9 Students: 1128, (593 E, 535 C) (4 cohorts)	6, 9	Students from 21 middle and high schools across multiple rural districts in Kentucky who were reading at least 2 years below grade level. 88% W, 26% SPED, 62% FRL.	**GRADE**	
						Grade 6	+0.08
								+0.10*
						Grade 9	+0.12*
SIM: Fusion Reading							
Schiller et al. (2012) Add'l period	SR/ARP	1 year	GRADE: Students: 581 (285 E, 296C)	6-10	Students from 7 schools (4 middle, 3 high) across 3 school districts in Michigan who scored between the 5th and 35th percentile on a reading screening test. 81% AA, 13% SPED.	**GRADE**	+0.05
						Comprehension	+0.08
						Vocabulary	0.00	+0.07
			MEAP: Students: 256 (118 E, 138 C)	6-7		MEAP Reading	+0.11

7/1/2019 11:30 AM

Table 9
Personalization Approaches

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Posttest	Effect sizes	Overall effect size
The Thinking Reader								
Drummond et al. (2011)	CR	1 year	Teachers: 90 (48 E, 42 C) Students: 2149 (1154 E, 986 C)	6	32 high-poverty schools from 16 districts in Connecticut, Massachusetts, and Rhode Island. 37% W, 28% H, 11% SPED, 10% ELL, 71% FRL.	GMRT	+0.01	+0.01
						Comprehension	+0.03	
						Vocabulary	-0.04	
Schoolwide Enrichment Model-Reading (SEM-R)								
Little et al. (2014)	CR	1 year	Teachers: 47 (27 E, 20 C) Students: 2028 (1198 E, 830 C)	6-8	4 high-poverty, low-achieving US middle schools. 48% AA, 32% H, 67% FRL.	GMRT Comprehension	+0.10	+0.10
Achieve3000								
Shannon & Grant (2015)	CR	1 year	Teachers: 33 (16 E, 17 C) Students: Grade 6: 494 (263 E, 231 C) Grade 9: 248 (122 E, 126 C)	6, 9	12 schools in 4 suburban and city districts across the US. 37% H, 67% W, 12% SPED, 12% ELL, 62% FRL.	GMRT	+0.29*	+0.29*
						Grade 6	+0.22	
						Grade 9	+0.44	
SuccessMaker								
Gatti (2011)	CR	1 year	Classes: 22 (11 E, 11 C) Students: 453 (254 E, 199 C)	7	Students reading at least 1 year below grade level from 5 schools in 8 urban and suburban school districts in 4 states (Arizona, Kansas, Michigan, Missouri). 51% W, 53% FRL.	GRADE	+0.11	+0.11
						Comprehension	+0.10	
						Vocabulary	+0.12	

Accelerated Reader

Gorard et al. (2015)	SR/TA	22 weeks	Students: 349 (166 E, 183 C)	Year 7	Low-achieving students in four schools in England. 88% W, 23% SEN, 35% FSM.	NGRT		+0.24* +0.24*

iLit

Gatti (2016)	SR/ARP	2 years	Students: 213 (114 E, 99 C)	7-8	Students reading below grade level from 6 middle schools in AZ, CA, CO, MI, NJ & NY. 53% H, 22% W, 17% AA, 26% LEP, 13% SPED, 80% FRL.	GRADE Total	+0.09	
						Comprehension	+0.12	+0.09
						Vocabulary	+0.01	

Table 10
Group/Personalisation Rotation Approaches

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Post-test	Effect sizes	Overall effect size
READ 180								
Schenck et al. (2011)	SR/ARP	2 years	Students: 1295 (556 E, 739 C) (3 cohorts)	6-8	Students from 8 Title I middle schools in Memphis City, Tennessee who tested in the bottom quartile of the reading/ELA portion of the state test. 93% AA, 6% ELL, 92% FRL.	ITBS Total Reading	+0.02	+0.02
						Comprehension	-0.01	
						Vocabulary	+0.06	
Meisch et al. (2011)	CR	3 years	Schools: 19 (10 E, 9C) Students 1023 (552 E, 471 C) (4 cohorts)	6-8	Students from 19 Title I middle schools across Newark, New Jersey who scored below proficient on state reading tests. 55% AA, 42% H, 13% ELL, 44% SPED, 62% FRL.	SAT-10		+0.06
						Comprehension	+0.06	
						Vocabulary	+0.05	
Swanlund et al. (2012)	SR/ARP	1 year	Students: 619 (335 E, 284 C)	6-9	Students from 5 Title I schools in Milwaukee who performed below proficient on standardised reading tests. 70% AA, 19% H, 36% SPED, 8% ELL, 88% FRL.	MAP Reading	+0.14*	+0.14*
Lang et al. (2009)	SR	1 year	High Risk Students: 190 (100 E, 90 C) Moderate Risk Students: 409 (207 E, 202 C)	9	Students from 7 comprehensive high schools in a large district in Florida who were reading below 4th grade levels (high risk) or between 4th and 6th grade levels (moderate risk). 19% H, 19% AA, 41% FRL.	FCAT Reading		+0.12
						High Risk	-0.27*	
						Moderate Risk	+0.30*	
Sprague et al. (2012)	SR/ARP	1 year	Students:456 (231 E, 225 C) (5 cohorts)	9	Students from 5 Title I eligible high schools in western Massachusetts who tested between a 4th and 6th grade reading level 73% minority, 19% SPED, 72% FRL.	SDRT-4	+0.18*	+0.18*

7/1/2019 11:30 AM

Expert 21 (Related to READ 180)

Study	Design	Duration	Sample	Grades	Setting	Measure	Effect	Overall
Sivin-Kachala & Bialo (2012)	CR/TA	1 year	Teachers: 6 (3 E, 3 C) Students: 276 (137 E, 139 C)	6-8	1 middle school in urban New Jersey. 71% H, 27% AA, 100% FRL.	NJASK	+0.20	+0.15
						Language & Literature	+0.22	
						Reading Comp.	+0.18	
						GMRT Comprehension	+0.10	

System 44 (Related to READ 180)

Study	Design	Duration	Sample	Grades	Setting	Measure	Effect	Overall
Beam et al. (2011)	SR/ARP/TA	1 year	Students: 147 (75 E, 72 C)	6-8	4 schools from one large suburban school district in southern California. 7% SPED, 63% W, 33% FRL.	TOSREC	-0.24*	-0.14
						CST	-0.04	
Beam & Faddis (2012)	SR/ARP/TA	1 year	Students: 145 (70 E, 75 C)	6-8	4 middle schools from one urban district in Michigan. 78% AA, 53% SPED, 96% FRL.	TOSREC	+0.20*	+0. 20*

Passport Journeys

Study	Design	Duration	Sample	Grades	Setting	Measure	Effect	Overall
Vaden-Kiernan et al. (2012)	SR/ARP/TA	1 year	iLEAP Students: 1102 (548 E, 554 C) GRADE Students: 983 (485 E, 498 C)	6,7	Students from 10 Title I middle schools across Louisiana who scored below proficient on state standardized reading assessments. 76% minority, 15% SPED, 88% FRL.	iLEAP Reading	-0.01	+0.12*
						GRADE Overall	+0.27*	
						Vocabulary	+0.13*	
						Comprehension	+0.30*	
Schenck et al. (2012)	SR/ARP/TA	1 year	SOL Students: 701 (343 E, 358 C) GMRT Students: 568 (279 E, 289 C)	7,8	Students from 9 middle schools in urban, high-poverty settings across Virginia who scored at least two years below grade level on reading tests. 68% AA, 24% SPED, 8% ELL, 88% FRL.	SOL Reading	+0.06	+0.06
						GMRT Overall	+0.06	
						Comprehension	+0.05	
						Vocabulary	+0.07	
Dimitrov et al. (2012)	SR/ARP/TA	1 year	Students: 460 (238 E, 222 C)	9	Students from 6 Title I high schools across Illinois who performed in the bottom two quartiles on the EXPLORE reading assessment. 58% AA, 5% H, 30% W, 18% SPED, 85% FRL.	GMRT	+0.02	-0.03
						EXPLORE	-0.09	

Comprehensive Circuit Training (CCT)

Fogarty et al. (2014)	CR/TA	1 semester	Classes: 61 (30 E, 31 C) Students: 859 (411 E, 448 C)	6-8	3 middle schools from 3 districts in the Southwest US. 43% H, 35% W, 22% AA, 9% ELL, 6% SPED, 67% FRL.	GMRT	+0.12	+0.12
Fogarty et al. (2016)	CR/TA	1 semester	Classes: 16 (9 E, 7 C) Students: 228 (112 E, 116 C)	6-8	3 middle schools in Texas. 30% AA, 27% W, 26% H, 9% SPED, 67% FRL.	STAAR +0.09 GMRT +0.12 GRADE +0.11 GORT +0.18	+0.15

Prentice Hall Literature

Eddy et al. (2010)	CR/TA	1 year	Teachers: 29 (16 E, 13 C) Students: 1518 (744 E, 774 C)	7-10	8 schools from California, Oregon, Arizona, Ohio. 6 suburban and 2 rural areas. 55% H, 15% AA.	GMRT	-0.10	-0.10

Strategic Adolescent Reading Intervention (STARI)

Kim et al. (2016)	SR/ARP	1 year	Schools: 8 Students: 398 (170 E, 228 C)	6-8	Schools in 2 urban, 2 rural districts in Massachusetts. Students were below 30th percentile on prior year MCAS. 49% W, 26% H, 19% AA, 13% ELL, 69% FRL.	RISE Reading Comp +0.08 Basic Reading +0.21* Vocabulary +0.16	+0.15

Read to Achieve

Deussen et al. (2012)	SR/ARP	1 year	Read to Achieve + PhonicsBlitz GMRT Students: 63 (32 E, 31 C) MSP Students: 76 (37 E, 39 C) Read to Achieve GMRT Students: 295 (144 E, 151 C) MSP Students: 325 (192 E, 191 C)	6, 8	Students from 5 Title I middle schools & 1 junior high school from 3 districts in Western Washington state who were reading at least two years below grade level. 43% W, 23% ELL, 58% FRL.	Read to Achieve + PhonicsBlitz GMRT +0.13 MSP +0.11 Read to Achieve GMRT +0.02 MSP +0.16	+0.10

Reading Intervention through Strategy Enhancement (RISE)								
Lang et al. (2009)	SR	1 year	High Risk Students: 194 (104 E, 90 C)	9	Students from 7 comprehensive high schools in a large district in Florida who were reading below 4th grade levels (high risk) or between 4th and 6th grade levels (moderate risk). 19% H, 19% AA, 43% FRL.	FCAT Reading		
						High risk	-0.06	+0.16*
			Moderate Risk Students: 406 (204 E, 202 C)			Moderate risk		+0.27*

Table 11
Intensive Approaches

Intervention	Design/ Treatment	Duration	N	Grades	Sample characteristics	Posttest	Effect sizes	Overall effect size
Reading Excellence: Word Attack and Rate Development Strategies (REWARDS)								
Newman & Kundert (2012)	SR/ARP	1 year	NYS: Students: 517 (253 E, 264 C) GMRT: Students: 469 (232 E, 237 C) (10 schools)	7	Students from 11 Title I middle schools across 4 boroughs of New York City who performed below proficient on state standardised reading test. 64% H, 22% AA, 12% A, 95% FRL.	NYS ELA	+0.15	+0.09
						GMRT Total	+0.02	
						Comprehension	-0.01	
						Vocabulary	+0.08	
Kentucky Cognitive Literacy Model (KCLM)								
Cantrell et al. (2012)	SR/ARP	1 year	Students: 485 (232 E, 253 C)	9	Students from 9 high schools in 9 districts who were reading at least two years below grade level. 88% W, 16% SPED, 62% FRL.	GRADE	-0.06	-0.06
REACH								
Lang et al. (2009)	SR	1 year	High Risk Students: 181 (91 E, 90 C) Moderate Risk Students: 401 (199 E, 202 C)	9	Students from 7 comprehensive high schools in a large district in Florida who were reading below 4th grade levels (high risk) or between 4th and 6th grade levels (moderate risk). 20% H, 20% AA, 43% FRL.	FCAT Reading		-0.02
						High risk	-0.19	
						Moderate risk	+0.06	

7/1/2019 11:30 AM