Published in final edited form as:

J Learn Disabil. 2017 ; 50(4): 450–467. doi:10.1177/0022219415618497.

Executive Function, Self-Regulated Learning, and Reading Comprehension: A Training Study

Paul T. Cirino, PhD[1], **Jeremy Miciak, PhD**[1], **Elyssa Gerst, MA**[1], **Marcia A. Barnes, PhD**[2], **Sharon Vaughn, PhD**[2], **Amanda Child, BS**[1], and **Emily Huston-Warren, BA**[1]

[1]University of Houston, TX, USA

[2]University of Texas at Austin, USA

Abstract

The goal of this study was to evaluate the extent to which training that emphasizes the process of executive function (EF) and self-regulated learning (SRL) would result in increased reading comprehension; we also evaluated interrelationships of EF, SRL, and reading. We report an experiment ($N = 75$ fourth graders) that contrasted two researcher-implemented conditions (text-based reading [TB] and text-based reading plus executive function [TB+EF]) to a control. We also evaluated relationships among measures of SRL, EF, and reading. Both the TB and TB+EF groups outperformed the control group for proximal text comprehension (where the topic was similar to that covered in training) and background knowledge related to it, but the two researcher-led groups performed similarly. There were no significant differences for less proximal text, and again similar performance for both TB and TB+EF. Correlations among measures were weak in general, although the pattern was similar to that found in the extant literature. The findings speak to the difficulty in separating these components from those of strong instruction more generally. The relationships of these constructs to reading comprehension will likely be enhanced by more sensitive measurement of EF and reading comprehension, particularly where tied to active treatment components.

Keywords

executive function; self-regulated learning; reading comprehension; training

Improving reading comprehension for struggling readers beyond the earliest grades has yielded a promising but evolving research base (Berkeley, Scruggs, & Mastropieri, 2010; Edmonds et al., 2009; Scammacca, Roberts, Vaughn, & Stuebing, 2013). However, effect sizes for intensive interventions with older students are generally less robust than those of younger readers, particularly on standardized measures of reading comprehension

Corresponding Author: Paul T. Cirino, Department of Psychology, University of Houston, 4849 Calhoun Rd., Suite 480, Houston, TX 77204-6022, USA. pcirino@uh.edu.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental Material
The online appendix is available at http://ldx.sagepub.com/supplemental.

(Scammacca et al., 2013; Scammacca, Fall, & Roberts, in press; Wanzek et al., 2013). One potential explanation for these differences is the change in reading task demands from a focus on more foundation skills early (e.g., word reading) to more advanced language and knowledge-based expectations as students read increasingly complex texts. For example, the overlap between word reading and comprehension is substantial in the earliest grades (e.g., K–Grade 2; Torgesen, 2000). As students move beyond beginning reading tasks and are confronted with more complex and cognitively demanding texts, the importance of vocabulary and listening comprehension increases (Catts, Hogan, & Adlof, 2005). Beyond the role of word decoding and listening comprehension in reading processes (i.e., the Simple View; Gough & Tunmer, 1986), inferencing and background knowledge have been linked to reading comprehension (e.g., the direct and inferential mediation model of reading comprehension; Cromley & Azevedo, 2007), although these factors may also be required for listening comprehension.

For students with word reading problems that were not adequately addressed early on, reading problems may compound over time. For other students, specific difficulty in reading comprehension may later emerge (Compton, Fuchs, Fuchs, Elleman, & Gilbert, 2008). For these students, factors less specifically tied to and beyond language-specific factors—namely, self-regulated learning (SRL) behaviors and executive functions (EFs)—may play a larger role in reading performance. Older struggling readers have experienced multiple years of reading failure and may exhibit corresponding deficits in motivation, negative attitudes toward reading, and other SRL-related weaknesses (Baird, Scott, Dearing, & Hamill, 2009; Logan, Medford, & Hughes, 2011; Sideridis, Morgan, Bostas, Padeliadu, & Fuchs, 2006; Solheim, 2011). SRL behaviors are important for reading comprehension because it is a goal-directed behavior (e.g., extracting meaning from text) that requires effort. To maximize knowledge gained from print, students must (a) engage with text to develop a specific goal (English, Barnes, Fletcher, Dennis, & Raghubar, 2010; Schunk, 1990), (b) associate text with their background knowledge (Elbro & Buch-Iversen, 2013; Pearson, Hansen, & Gordon, 1979), (c) make inferences within text (Cain & Oakhill, 1999; McGee & Johnson, 2003), and (d) evaluate their reading performance to ascertain whether their performance is adequate (Schunk, 2003).

There is considerable research linking EFs to reading comprehension (Best, Miller, & Naglieri, 2011; Cartwright, 2012; Christopher et al., 2012; Locascio, Mahone, Eason, & Cutting, 2010; Sesma, Mahone, Levine, Eason, & Cutting, 2009). EF is commonly operationalized as measures of cognitive functions (e.g., shifting, inhibition, working memory, planning, fluency; Jurado & Rosselli, 2007). Like SRL, EFs are associated with goal-directed behaviors (Lezak, Howieson, Bigler, & Tranel, 2012) and may be viewed from a problem-solving framework by which goals are set, plans/strategies are developed, performance is executed and monitored, and performance is evaluated (Zelazo, Carter, Reznick, & Frye, 1997). Through a connection to goal-directed behavior, the construct of EF shares much in common with the construct of SRL (Zhou, Chen, & Main, 2012). Indeed, some researchers (Ilkowska & Engle, 2010; Ylvisaker & Feeney, 2002) have made reference to a combined SRL/EF construct.

Given the hypothesized role of both EF and SRL to the reading process, specific training that incorporates SRL and EF principles may hold potential as an adjuvant intervention in the context of more typical reading intervention approaches, particularly for readers as they focus more on comprehension than word reading. The purpose of this article is to evaluate the role of EF/SRL to reading comprehension within the context of an experimental study of struggling readers in Grade 4. We do so by experimentally comparing training conditions that differ in their incorporation of EF/SRL principles and by examining the relationship of aspects of EF and SRL to reading comprehension performance.

EFs and Reading Comprehension

At least three aspects of EF can be viewed as particularly relevant to reading comprehension: planning, fluency, and working memory. Correlationally, measures of planning are uniquely related to reading performance (Cutting, Materek, Cole, Levine, & Mahone, 2009; Sesma et al., 2009) although not always (Locascio et al., 2010; Sikora, Haley, Edwards, & Butler, 2002). Measures of planning (e.g., Tower of London; Shallice, 1982) are informative because their successful completion involves preparing for action and monitoring progress to solve a problem. Generative fluency, as indexed by measures of rapid lexical (e.g., generate words beginning with a letter) or semantic (e.g., generate words belonging to a category) retrieval, is germane in that it may be reflective of a student's ability to generate and associate background knowledge to the content of to-be-read text. Fluency in a more general sense also may be important to generate relevant questions and dialogue related to text monitoring during the reading process and to make reflective statements after reading text. However, no known studies specifically evaluate the role of verbal fluency in reading comprehension in struggling readers, but such measures are found in group-based analyses (e.g., Cohen, Morgan, Vaughn, Riccio, & Hall, 1999; Stern & Morris, 2013). Measures of verbal fluency also may reflect working memory (see below) given that restriction rules must be adhered to and words cannot be repeated.

There is strong support for the role of working memory in reading comprehension (Cain, Oakhill, & Bryant, 2004; Carretti, Borella, Cornoldi, & De Beni, 2009; Savage, Cornish, Manly, & Hollis, 2006; St. Clair-Thompson & Gathercole, 2006). Although predictive relationships with reading are common, experimental manipulations of working memory have also been explored. In these studies, students are randomized to a working memory training group or to control (CON) and are compared on trained tasks and/or near- and far-transfer tasks (the latter of which can include academic skills). However, evidence that such interventions improve reading comprehension is elusive (for reviews, see Melby-Lervag & Hulme, 2013; Shipstead, Redick, & Engle, 2012). Therefore, if the goal is improvement in reading comprehension (as in the present study), direct cognitive training may yield less robust results than if EF principles are integrated within explicit reading instruction.

SRL and Reading Comprehension

SRL can be defined as "processes whereby learners personally activate and sustain cognitions, affects, and behaviors that are systematically oriented toward the attainment of personal goals" (Zimmerman & Schunk, 2011, p. 1). Specific components of SRL linked to reading include (a) the use (activation) of background knowledge (Braten, Ferguson, Anmarkrud, & Strosmo, 2013; Taboada, Tonks, Wigfield, & Guthrie, 2009), (b) strategy use (Dignath, Buettner, & Langfeldt, 2008; Diseth, 2011; Stroud, 2006), (c) self-efficacy (Caprara et al., 2008; Shell, Murphy, & Bruning, 1989; Solheim, 2011; Zuffiano et al., 2013), (d) motivation (Guthrie, Wigfield, Metsala, & Cox, 1999; Wolters, Denton, York, & Francis, 2014), and (e) performance goal orientations (Diseth, 2011; Hornstra, van der Veen, Peetsma, & Volman, 2013). However, there is variability in both the size and the consistency with which these relationships exist, and correlations less than $r = .30$ are common.

Intervention studies that manipulate SRL components as a vehicle for improving learning may have as their outcome the SRL components themselves (e.g., increased self-efficacy or motivation or change in attributional stance; e.g., Nelson & Manset-Williamson, 2006). More relevant to the present work, studies manipulating SRL may also include reading outcomes such as reading comprehension. Studies that do so often focus on a specific instructional practice (e.g., using graphic organizers, question asking, testing and spacing effects, summarization, mnemonics); some of these have been summarized in Institute of Education Sciences practice guides on effective practices for reading (Kamil et al., 2008; Shanahan et al., 2010).

A number of brief training studies using SRL principles have also shown effectiveness in improving reading comprehension (e.g., Antoniou & Souvignier, 2007; Berkeley, Mastropieri, & Scruggs, 2011; Mason, 2013; Miranda, Villaescusa, & Vidal-Abraca, 1997; McGee & Johnson, 2003; Zentall & Lee, 2012). For example, within a broader review of self-regulated strategy development, Mason (2013) reviewed two studies demonstrating the effectiveness of the three-step “TWA” process (Think before reading, think While reading, think After reading) for reading comprehension. In one study, on an outcome measure of oral retelling, students in TWA significantly outperformed those in a reciprocal questioning condition; in another study, students in the TWA condition significantly outperformed CON group members on three comprehension measures (a guided reading group did not differ significantly from the other two conditions, although effect sizes favored TWA). The division of activities into prior, during, and after a given event overlaps with the phases of models of SRL (e.g., Zimmerman, 2000), and this is the overall framework within which the EF and SRL components were added to the text-based treatment condition in the present experiments.

The Present Work

As noted previously, there are commonalities between EF and SRL, though there are also challenges in relating them to one another. Both of these constructs are conceptualized in multiple ways, with findings reported for each construct in distinct literatures: SRL is typically represented in education-related journals and EF in neuropsychology-related journals. That both EF and SRL are implicated in reading comprehension amplifies the connections between them. One potential difficulty is being able to extract EF/SRL specific components from good teaching/instructional practice, as motivation and engagement are often included as key components of effective literacy instruction and intervention (e.g., Kamil et al., 2008).

The purpose of the current work is to evaluate the impact of EF/SRL on reading comprehension using a multifaceted approach to intervention and assessment in the context of a training experiment. We construed EF/SRL broadly, including (a) neuropsychological performance measures of EFs (e.g., planning, fluency), (b) measures of SRL processes (e.g., effort/self-efficacy), and (c) instructional guidance that addressed components of successful goal-directed behavior (e.g., goal setting, planning, implementing, monitoring, evaluating) around reading comprehension.

Two researcher-implemented treatments, one focused on text (text-based reading [TB]) and another focused on text that additionally integrates principles associated with EF/SRL into the text-based strategies (text-based reading plus executive function [TB+EF]), are contrasted with one another and with a CON group. We also report correlations of measures of SRL and EF to one another and to reading. To the extent that SRL and EF both reflect goal setting, generation, and monitoring/reflection, measures of each could be expected to relate to one another, and to the extent that such a process is important for reading comprehension, each should also demonstrate relationships with this skill.

The primary goal of this study was to evaluate the extent to which training that emphasized the process of EF would result in increased reading comprehension. We hypothesized that students randomized to the TB+EF condition would outperform students in the TB condition, who in turn would outperform students randomized to a business-as-usual CON condition, for proximal text (the same era of state history). We also expected this advantage to extend to related text in the same domain (another era of state history) and to informational text reading in an unrelated domain (e.g., science of forces and motion). A second goal was to relate measures of EF and SRL to one another and to measures of reading. Based on prior literature, we expected these skills to relate to reading moderately, and to one another more modestly.

Method

We first briefly describe a pilot procedure that helped inform the parameters of implementation for the experiment. The pilot included 31 third-grade students (age = 9.67 years; $SD = 0.73$) who were randomized to one of three study conditions (9 TB, 10 TB+EF, 12 CON). There were two tutors, and each taught both conditions in groups of 5 students for 10 daily intervention sessions of approximately 35 min each during 2 weeks. The two conditions covered the same content (state history) and were equated for amount of material, time reading aloud, and overall instructional time. The activities for both active conditions (for both the pilot and the primary experiment) appear in Table 1.

Groups did not differ on the pretest (proximal text for to-be-covered material), $F(2, 28) = 3.21$, $p = .055$. However, effect sizes were sizable (TB+EF to CON, $d = -1.21$; TB to CON, $d = -0.36$; TB+EF to TB, $d = -0.54$); in other words, the TB+EF group performed well below the levels of the other conditions. Groups did not differ on standardized reading measures. For proximal text at posttest (with pretest as covariate), the overall model was significant, $F(3, 27) = 5.74$, $p = .004$. The pretest variable was significant ($p < .001$), but the treatment effect was not, $F(2, 27) = 3.20$, $p = .056$. The effect size for the raw score comparison of TB+EF to CON was $d = +0.09$, and for TB to CON it was $d = +0.60$; corresponding effect sizes constructed from least squares means were +0.85 and +1.03, respectively (greater improvement in the active conditions). Quiz performance of the TB+EF group was also evaluated; the first five quizzes were averaged and then compared to the remaining six (roughly Week 1 vs. Week 2). The within-groups effect was significant, $F(1, 9) = 41.81$, $p < .001$; during the first five lessons, students answered $M = 2.47$ questions correctly (out of 5), and for the remaining six lessons, $M = 3.67$. Analysis of other texts yielded a similar directional pattern of results. For example, for near transfer text, the effect size for the raw score comparison of TB+EF to CON was $d = +.11$ and for TB to CON was $d = +.23$ (with effect sizes constructed from least squares means of +0.77 for the TB+EF to CON difference and +0.58 for the TB to CON difference).

The findings of the pilot study guided us in modifying some aspects of the treatment and in conducting the target experiment. Specifically, the TB+EF condition had a written "check-in" activity that took too much time and did not allow for discussion. Also, there were too many activities in the TB+EF condition to enact fully within the time frame. Further structure around behavior and motivation might have enhanced results as well, especially given that the training took place at the end of the academic year. The brevity of the training and the fact that the students were not explicitly identified as struggling readers were additional factors that may have occluded the ability to systematically distinguish between the conditions. These issues were addressed in the target study.

Participants

A total of 252 students in Grade 4 from three schools (different from those in the pilot procedure) in a southwestern metropolitan area were screened on a measure of reading comprehension (*Gates-MacGinitie*; MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2000). Of these, 84 (33%) met the participation criterion of a standard score of 90 or less. This percentage reflects the higher risk rate of students attending these schools, although all schools were rated as academically acceptable by the state accreditation agency. The present report focuses on 75 students who completed the study. Table 2 provides demographic information by group.

Procedures

Students were randomized within school to one of the three study conditions (TB, TB+EF, CON). Of the 84 students who met initial eligibility criteria, 3 were not randomized (2 due to limited English proficiency and 1 who withdrew). Of the 81 students who were randomized, 1 student withdrew after randomization, and 5 additional students were excluded due to scheduling/resource issues (after they had been randomized to one of the two active treatments). One student was not pretested due to examiner error. Schools varied in the times of day for sessions as well as in the length and maximum number of sessions. However, these parameters were equivalent across groups (see Table 5 later in this article). Posttest occurred the week following the training.

Experimental Conditions

There were three conditions (same as in the pilot study) to which students were randomly assigned (TB, TB+EF, and CON). The topic (state history) remained the same, and the amount of text read at any one time was in sections of approximately 75 to 150 words. Every condition was delivered in each school. Descriptions follow, and supplemental materials provide further details (http://ldx.sagepub.com/supplemental).

TB condition—On the first day, the structure and sequence of the sessions were previewed, as were explicit behavioral expectations and a reward system. The latter were set up as a points system, with students' gaining points for specific fixed events (e.g., appropriate entry, silent/oral reading) as well as for appropriate behavior at random times (Surprise Time! denoted by a small random timer twice per session). Points were exchanged for small prizes at regular intervals. In both conditions, feedback was provided to students frequently in a manner that was immediate and specific. Also on the first day for both groups, three short, unrelated passages were used to demonstrate the structure of the lessons. For the first passage, the tutor led the group through the passage, talking through each step; for the second passage, the tutor solicited "help" from students in the group; and for the third passage, the tutor encouraged students to perform all steps, with help as needed.

After the first day, the TB condition consisted of six activities (see Table 1), several of which were iterated for every text passage. After each lesson, students in the TB group received a four-item multiple-choice quiz covering the lesson content, with each answer identifiable directly from text.

TB+EF condition—As with the TB condition, the first day for the TB+EF condition included a review of the study purpose, group behavioral expectations, and reward system. Elements related to SRL and EF were emphasized, which included further specific discussion around motivation as well as highlighting of activities to be conducted prior to reading, during reading, and after reading. Activities for the TB+EF condition appear in Table 1. After each lesson, students in the TB+EF condition also received a four-item multiple-choice quiz covering lesson content. Two questions overlapped with those of the TB condition: one asked about a covered vocabulary word, and one asked about the main idea. Because most students in both conditions missed one or more days of tutoring, and because groups completed a slightly different total number of lessons, quiz data were analyzed as average quiz score across a given week comprising the dependent variable.

Tutors and Fidelity

There were five tutors. Each of the tutors was an experienced teacher and/or interventionist, and three of five held elementary teaching credentials. Four of the tutors taught both active conditions, and the fifth tutor taught only one (TB+EF) group. All tutors received approximately 8 hr of training prior to implementation of the intervention. In addition, two observers (authors) visited all schools at least weekly—and each tutor at least four times—to measure intervention fidelity and provide immediate feedback to tutors as well as more informal coaching on an ongoing basis. The fidelity instrument rated implementation of the intervention within each intervention component for the two conditions. Thus, the form differed by condition, with eight items for the TB+EF condition and six items for the TB condition reflecting the number of instructional components for each condition. A total of 34 fidelity observations were coded. For each component, implementation was assigned a score on a Likert-type scale from 0 (*poor fidelity/not implemented*) to 4 (*highest fidelity*). Across components, mean fidelity scores were good (M = 3.16, SD = .36, range = 2.5–3.7), indicating that each intervention component was implemented in the good to very good range, suggesting overall compliance with the treatment. Tutors were also rated on a Likert-type scale from 1 to 5 for overall fidelity in program implementation (M = 4.02, SD = .74) and quality of instruction (M = 4.11, SD = .68). All fidelity observations were treated as opportunities for on-site coaching.

Measures[1]

Standardized descriptive measures—The screening measure was the *Gates-MacGinitie Reading Tests* (4th ed.; MacGinitie et al., 2000); it is commonly used and demonstrates good psychometric properties (reliability coefficients for reading comprehension exceed .90). Students were also administered the *Woodcock-Johnson III Test of Achievement* (WJ-III; Woodcock, McGrew, & Mather, 2001) Letter-Word Identification subtest, which assesses the ability to read real words. Reliabilities exceed .90. Reading fluency was assessed via the *Test of Word Reading Efficiency* (Torgesen, Wagner, & Rashotte, 1999) Sight Word Efficiency subtest, for which students are given a list of 104 words and are asked to read them as accurately and as quickly as possible within 45 s. Psychometric properties are good, with most alternate forms and test-retest reliability coefficients greater than or equal to .90 in this age range. The *Test of Silent Reading Efficiency and Comprehension* (Wagner, Torgesen, Rashotte, & Pearson, 2009) is a timed sentence verification task. Corrected alternate-form reliability coefficients for students in Grade 4 range from .82 to .86. Finally, reading comprehension was assessed with the WJ-III Passage Comprehension subtest. This is a cloze-based reading test in which students are required to supply a missing word from a sentence or longer passage that they read silently. Reliability coefficients are good, ranging from .80 to .92 for students ages 8 to 12 years.

Experimental pre-post measures—There were three direct text outcomes. Passages ranged from 432 to 472 words and contained expository content. The lexile rating of the passages ranged from 840L to 1000L, consistent with the band of lexiles for Grades 4 to 5 associated with Common Core content (740L to 1010L; downloaded from www.lexile.com). Students had 15 min to read each passage silently and answer 12 multiple-choice questions about them. One was proximal text (the content but not actual text was covered in the training), which was administered at pretest and posttest, and two additional texts were administered only at posttest (one related to taught content—a different era of state history and an unrelated text for science content—forces and matter). Pilot procedures helped establish the content of and multiple-choice questions for the text outcomes; in that pilot, α = .66 for proximal text at pretest and .70 at posttest. Related text showed α = .64, and unrelated text showed α = .71. Several students did not respond to all questions within the

[1]Further details about the measures used here, and in the parent project as a whole, can be found at http://www.texasldcenter.org/projects/measures.

time limit; for consistency, the dependent measure was the percentage of attempted items answered correctly. The proximal text at pretest was the covariate for all text outcomes. Across time points, in the sample as a whole, the standardized and experimental reading tasks correlated moderately with one another (range r = .07–.45, median = .21). The experimental reading tasks also correlated moderately among themselves (range r = .26–.40, median = .36), as did the standardized measures (range r = .12–.59, median = .29).

In the Background Knowledge task, students read silently and answered 10 multiple-choice questions on state history to establish the extent of domain knowledge prior to tutoring. This measure was given at pretest and posttest, in a group format; 15 min were allotted.

SRL measures—In models of SRL (e.g., Zimmerman, 2000), tasks unfold over three overlapping phases:

1. those prior to task initiation (Forethought), emphasizing goal setting, planning, and preparedness;
2. those during the task (Performance), emphasizing monitoring and control; and
3. those that occur with task completion (Reflection), emphasizing recall and evaluation.

In addition, the roles of motivation, self-efficacy, and goal directedness are important in such phases (e.g., Pintrich, 1999, 2000). Measures relevant to SRL are described below. Some are designed to explicitly measure specific phases of SRL (Titles, How Will I Do? [HWD], How Did I Do? [HDD]); others measure beliefs important for SRL (*Student Learning Questionnaire* [SLQ]), and still others assess behavioral self-regulatory skills (*Strengths and Weaknesses of ADHD Symptoms and Normal Behavior* [SWAN]; Swanson et al., 1998; Swanson et al., 2012).

In the Titles task (a measure of forethought), students are asked to set goals for reading, based on a title. Students are shown cards with titles and are asked to state what they know about the topic (background knowledge) and what they might find out from reading a passage with that title (goal setting). Students are shown five such titles, with responses audio-recorded to eliminate writing burden. The task required about 10 min and was given only at posttest.

For the HWD task (a measure of forethought), students mark on a line corresponding to their estimate of their performance on subsequent reading tasks. There are six “questions” in this vein that are read aloud as students follow along (“I am going to try my best …”; “I am ready to pay attention …”; “I am going to think about what I know …”; “I am going to set goals …”; “I am going to ‘check-in’ with myself …”; “I am confident I will get many answers correct …”). One measure of reflection was the HDD task, which is analogous to HWD, but questions are phrased in the past tense (e.g., “I tried my best …”); it was given following the reading measures (as a measure of reflection). Each task requires approximately 3 min, and each was given at both pretest and posttest. Coefficient alphas of these measures were α = .57 (HWD) and α = .71 (HDD); average item correlations were r = .31 and .45, respectively.

Elements of the SLQ (available from first author) were utilized. The SLQ was developed to assess elements of SRL and reading in particular. The full SLQ has 42 items arranged onto four subscales supported by latent variable analysis. Students in this study received a briefer version (28 items), although prior to the aforementioned psychometric study; as a result, only items from the Effort/Self-Efficacy (of Learning) subscale (13 items) and the Perceived Skill/Preference (for Reading) subscale (5 items) are considered here. Students are read items orally, and they answer by marking a bubble-sheet on a scale of 0 (*not like me*) to 3 (*very much like me*). In a much larger sample of similar age (N = 896 students), the Effort/Self-Efficacy subscale showed α = .84, and the Perceived Skill/Preference scale showed α = .66. These subscales were administered at both pretest and posttest.

The SWAN rating scale was utilized at posttest only. The SWAN is an 18-item questionnaire based on the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, diagnostic criteria for attention-deficit/hyperactivity disorder (Swanson et al., 1998; Swanson et al., 2012). Teachers in this study rated the students on a 7-point scale, with lower scores indicating greater problems with attention. Continuous scales of Inattention and Hyperactive/Impulsive were included in the analyses.

EF measures—The Tower Task (after Shallice, 1982) is a planning task wherein rings or balls are loaded onto one of three sticks. Students see an initial configuration and a target configuration, and they must make their model match the target in as few moves as possible while obeying several rules. This task is computer based and uses the Inquisit platform, which is software for displaying, editing, and performing psychological experiments. In a similar sample (N = 839), the reliability coefficient of this measure was α = .69. Given the planning nature of this task, it was expected to relate to SRL measures of forethought. The *Delis-Kaplan Executive Function System* (D-KEFS; Delis, Kaplan, & Kramer, 2001) Verbal Fluency subtest has three conditions in which students have 1 min to orally produce as many exemplars of a target as possible (conditions include words that begin with each of three letters, two categories, and then words that switch between two categories). All task parts require approximately 9 min. Reliability coefficients range from .53 to .70 for students ages 8 to 19 years (Delis et al., 2001). Because this measure requires both generativity and retrieval, it was expected to relate to measures of both forethought and reflection. All cognitive measures were administered only at pretest.

Analysis Plan

The study design was a randomized pre-post design; we utilized ANCOVA for primary outcomes, in line with recommendations from reviews of their strengths and weaknesses relative to other approaches such as repeated measures or posttest ANOVA (Dimitrov & Rumrill, 2003). The covariate used for all reading outcomes (proximal, related, unrelated) was pretest reading, and school was also explicitly included in these models. We also computed effect sizes (Cohen's d) on the three primary posttest measures and then corrected for bias. Other analyses for group differences on self-regulatory measures (SLQ, HWD/HDD, Titles) used a mixed repeated measures analysis (with treatment condition as a between-groups factor and other within-group factors, depending on the measure). Correlations established relationships between reading, EF, and SRL measures.

Results

Pretest

Across all measures administered at pretest (background, pretest proximal text, standardized reading variables, SRL tasks, performance measures of EF, rating scales), there was only one between-groups difference (with CON students rating themselves as doing better than TB +EF on the HDD task; all other differences $p > .05$). For example, on the proximal text pretest variable, the overall model was not significant, $F(2, 71) < 1$. This pretest variable was used as the covariate for all direct reading comprehension outcomes.

Group and overall means and standard deviations appear in (a) Table 3 for descriptive measures, (b) Table 4 for treatment-related measures (e.g., pretest and posttests), (c) Table 5 for SRL measures, and (d) Table 6 for neuropsychological EF measures.

Proximal material—First, we present information for outcomes most proximal to the content of the training (proximal text, quizzes obtained during training, background knowledge). Performances on all three posttest text measures are provided in Table 4, by condition. For the text most proximal to taught content, including school and pretest, the overall model was significant, $F(5, 68) = 5.72$, $p < .001$. The pretest variable was significant ($p < .001$), as was the treatment effect, $F(2, 68) = 5.16$, $p = .008$; school was noncontributory ($p > .05$). Raw score effect sizes for TB+EF to CON was $d = +0.42$ and for TB to CON was $d = +.71$; effect sizes constructed from least-squares means were similar although slightly larger ($d = +0.60$ and $+0.89$, respectively). The raw effect size difference between TB+EF to TB was $d = −0.26$.

On the lesson quizzes (mean performance across each of the 4 weeks of tutoring), for the two shared items (where conditions could be directly compared), there was no effect of group, $F(1, 40) < 1$. There was also no Group × Time interaction, $F(3, 120) = 1.06$, $p > .05$. The effect for time was significant, $F(3, 120) = 17.30$, $p < .001$. Follow-up revealed an uneven pattern with high initial performance (Lessons 1–4, $M = 1.52$) that declined (Lessons 5–8, $M = 1.07$) before improving (Lessons 9–12, $M = 1.45$; Lessons 13–16, $M = 1.56$); a similar pattern was evident on the full quizzes within each group.

For background knowledge, with pretest performance of this measure included, the overall model was significant, $F(5, 68) = 6.53$, $p < .001$. The pretest variable was significant ($p < .002$), as was the treatment effect, $F(2, 68) = 11.80$, $p < .001$; school was noncontributory ($p > .05$). Effect sizes for TB+EF to CON was $d = +1.09$ and for TB to CON was $d = +0.94$; for TB+EF to TB, it was $d = +0.10$.

Results for related text yielded an overall model that was significant, $F(5, 68) = 2.81$, $p = .023$. The pretest (proximal text) variable was significant ($p = .025$), but the treatment effect was not, $F(2, 68) = 1.42$, $p = .249$. Effect size for TB+EF to CON was $d = +0.23$; for TB to CON, $d = +.42$; and for TB+EF to TB, $d = −0.23$; least-squares effect sizes were similar. For unrelated (science) text, the overall model was significant, $F(5, 68) = 2.86$, $p = .021$; the pretest (proximal text) variable was significant ($p = .004$), but the treatment effect was not,

$F(2, 68) < 1$, and effect sizes were negligible (TB+EF vs. CON, $d = +0.05$; TB vs. CON, $d = -0.08$; TB+EF vs. TB, $d = +0.10$).

SRL measures—Means and standard deviations for these measures, by group, are included in Table 5. Three repeated measures analyses were used to evaluate group differences on these measures. For the SLQ (self-report of learning-related behaviors), the between-groups variable was treatment condition, and the within-groups factors were time (pre and post) and type (the effort/efficacy and skill/preference subscales). The between-groups factor did not reveal overall group differences, $F(2, 68) < 1$. There was an effect for type ($p < .0001$), although the effort/efficacy subscale had higher scores merely because it had more questions. There was no effect for time ($p = .225$), and there were no interactive effects of time with treatment ($p = .507$) or time with type ($p = .415$); in addition, the Time × Type interaction ($p = .410$) and the Time × Type × Treatment interactions ($p = .726$) were not significant. Thus, there were no statistically significant differences between conditions on these scales over time or as they related to treatment.

For the HWD/HDD tasks, the between-groups variable was treatment condition, and the within-groups factors were time (pre and post) and task (the HWD task and the HDD task). The between-groups factor did not show an overall group difference, $F(2, 70) = 2.70$, $p = .074$. There was no effect for task ($p = .671$), although there was an effect for time, $F(1, 70) = 7.70$, $p < .001$, with confidence improving over time (72% at pre; 76% at post). There were no two-way interactive effects of time with treatment ($p = .518$), time with task ($p = .684$), or time by task ($p = .139$). However, the Time × Task × Treatment interaction was significant, $F(2, 70) = 4.48$, $p < .015$. Follow-up of this three-way interaction revealed that for the CON group members, their confidence increased for prereading performance (HWD; from 71% to 81%) pre- to posttraining, but postreading perceptions of performance did not change (HDD; 79% to 78%). For TB, there was no Time × Task effect. For TB+EF, there was a larger increase in confidence for HDD pre- to posttraining (from 65% to 73%), relative to HWD (70% to 72%).

For the Titles task, the between-groups variable was treatment condition, and the within-groups factors were question ("What do you know?" vs. "What might you find out?") and type (number of items stated; number of items stated that were reasonable/accurate). Recordings of 12 students (7 in CON group, 3 in TB group, and 2 in TB+EF group) were not useable due to examiner error. The between-groups factor did not show an overall group difference, $F(2, 60) = 1.12$, $p = .334$. There was an effect for question, $F(1, 60) = 6.33$, $p = .015$, with the first question (What do you know?) generating more responses. There was also an effect for type, $F(1, 60) = 111.33$, $p < .001$, with more stated than reasonable items, as would be expected. There was also a question by type interaction, $F(1, 60) = 64.12$, $p < .001$, with the gap between stated/reasonable responses greater for the first question ($M = 13.32$ to $M = 8.10$) than for the second question ($M = 10.40$ to $M = 8.46$). However, there were no interactions of treatment with question ($p = .901$), with type ($p = .130$), or with question and type ($p = .585$).

Correlations—Means and standard deviations, by group, for the neuropsychological measures of EF are included in Table 6. In the sample as a whole at pretest, experimental and standardized reading measures correlated with the two EF tasks modestly (range $r = |.01–.40|$, median = .14), with the strongest relationships between the two standardized reading comprehension measures and D-KEFS Verbal Fluency (letters and categories) and Tower Task (excess moves; range $r = |.23–.40|$). Results were similar for the relationship of SRL tasks with reading (range $r = |.01–.29|$, median = .12), with HWD/HDD variables correlated with the experimental pretest measure and WJ-III Word Identification (rs = .20 and .29, respectively). In general, cognitive EF tasks (at pretest) were weakly related to the experimental reading posttest measures, although Tower Task excess moves were negatively related to these (range $r = -.17–.33$). There were also weak relationships among SRL posttest measures with experimental reading posttests (range $r = |.01–.33|$, median .09).

Correlations among EF (at pretest) and SRL measures (at both pretest and posttest) appear in Table 7. As shown in the table, these variables were weakly related to one another (range $r = |.01–.33|$, median = .10). D-KEFS Verbal Fluency Switching had the most consistent relationship with other measures, including HWD/HDD measures at pretest and posttest, SWAN ratings (posttest), and the skill/preference subscale (pretest; range $r = |.16–.33|$, median = .23).

Discussion

This study examined the relationship of EF, broadly construed, to reading comprehension, in the context of a training study. The primary hypothesis was that the instructional groups with a focus on text plus an added EF component would outperform a CON group as well as a group receiving reading instruction with a focus on text. The primary results indicated that students learned the content of the material to which they were exposed, in terms of reading comprehension as well as background knowledge; however, the improvement was not differential in relation to TB instruction. There was little transfer to less proximal (related) text ($d = +0.23$ TB+EF to CON; $d = +0.42$ TB to CON), and none for unrelated (science) text, again with little difference between the two researcher-implemented trainings. The secondary hypothesis was that measures of reading would relate to measures of both SRL and neuropsychological measures of EF. Results here indicated generally weak relationships of reading performances with self-regulation/EF measures and of the latter with one another.

Training Study

In the experiment (and the preceding pilot), students improved their comprehension of the content to which they were exposed. Students in both of the active experiment conditions increased their performance on reading measures and background knowledge of this material. That the TB+EF group showed improvement for the material covered is encouraging, though the lack of appreciable differences relative to the TB group complicates findings. However, this type of finding is not unusual in the literature. For example, Johnson, Graham, and Harris (1997) had four small active training groups of students with reading disability during a brief period and found broad improvements for all four groups from pre to post. While strategy instruction improved reading comprehension performance, no added benefit was noted when components of goal setting (practice with setting, implementing, and monitoring goals) and/or self-improvement (using self-statements to guide strategies) were included. More recently, Spörer and Schünemann (2014), in a large-scale study of students with a range of skill, studied four active groups—all included reciprocal teaching, and three also included self-regulatory (of strategy implementation and/or outcome regulation) components. For reading comprehension outcomes, effect sizes at posttest for the groups with self-regulatory components were small relative to reciprocal teaching (g = .10–.16) though larger at maintenance (.27–.43). Effect sizes for reading motivation outcomes were moderate for the individual self-regulatory groups relative to reciprocal teaching alone (.47 and .48), though interestingly not for the combined self-regulatory group. Berkeley et al. (2011) found gains relative to their CON (read naturally) for both their reading strategy condition, and a reading strategy condition that added attributional training. Effect sizes were large for both a measure of written summarization and a questionnaire of specific strategy use, though the three conditions did not differ on a constructed response and multiple-choice passage test. Attributions for success (though not failure) increased for the attributional condition relative to controls at posttest and delay (with large effect sizes).

There are at least three key potential reasons that the TB+EF group did not show differential improvement relative to the TB group. These reasons include (a) overlap among the two active conditions (particularly behavioral/motivational components), (b) the effects of the TB+EF condition incorporating many "moving parts," and (c) challenges in measurement of the targeted constructs. Each is discussed in turn.

Overlap—The TB and TB+EF conditions shared numerous components, resulting in a strong TB contrast condition. Johnson et al. (1997) described their contrast condition similarly. The fact that both TB and TB+EF groups showed improvement relative to controls in this and other studies supports this interpretation. As shown in Table 1, the primary unique components of the TB+EF condition focused on preparation to read (e.g., titles, goals, get ready) and monitoring (e.g., check-in), whereas the primary unique component of TB was answering questions based on text. However, both active treatment conditions included (a) motivational aspects (feedback, reward system); (b) tutor-guided oral reading and additional silent reading for every lesson text portion; (c) regular quizzing of recently read material, which has been supported as a strong learning technique (e.g., Karpicke & Roediger, 2006; Rawson & Dunlosky, 2013; Roediger, Putnam, & Smith, 2011); and (d) summarization; both active conditions were also (e) equated for time with text and overall duration. Thus, the failure to find robust differences between researcher-implemented groups may reflect the efficacy of shared intervention components rather than the ineffectiveness of components specific to either of the two active treatment conditions. Two in particular stand out beyond the effects testing and other specific cognitive science–informed learning techniques—rereading and motivation.

First, there is considerable evidence to support the positive effects of repeated reading as a strategy to improve reading fluency and comprehension (Chard, Vaughn, & Tyler, 2002; Kuhn & Stahl, 2000; Therrien, 2004). The National Reading Panel (2000) also found that repeated reading led to consistent and positive gains in word recognition, fluency, and comprehension. By including a robust component conceptually similar to repeated reading in both treatment conditions, we may have unintentionally obscured our ability to identify the presumably smaller effects of intervention components specific to either of the active conditions.

Second, the fact that both the TB and the TB+EF conditions included behavioral scaffolds and motivational components may have affected results. Motivation and engagement have been demonstrated to improve reading; for example, these are reviewed as showing moderate support for reading for both younger (K–3 students; Shanahan et al., 2010) and older readers (adolescents; Kamil et al., 2008). The fact that such components are important for not only reading but also math (e.g., Fuchs et al., 2008) demonstrates the potential generalizability of such effects. Motivation is often considered "self-regulatory" in nature, and it could be argued that stronger differential effects would have manifested if only the TB +EF group had access to these scaffolds. Whether or not this were the case, these aspects are used quite widely in most teaching, especially intervention contexts, and are effective in that regard. Therefore, we deemed it necessary to include motivational and behavioral components in both of our active conditions. Such issues speak to the difficulty of disentangling what constitutes an effective "executive" or "self-regulated learning" component from practices supported by cognitive science (e.g., strategies, techniques associated with improved learning) or from good instructional techniques more generally. It may still be the case that some types of motivational and/or behavioral techniques are more effective than others, and future studies designed to test this question directly would be of value.

Many components—The fact that the TB+EF condition had a larger number of activities than the TB condition also may have inadvertently diluted differential effects, potentially due to cognitive "overload." A similar suggestion was made by Spörer and Schünemann (2014) for their study. This is despite the fact that one change made from the pilot procedure to the current experiment was to decrease the number of activities in each TB+EF session. Still, given the relatively greater complexity of the TB+EF condition (relative to that of TB), it is perhaps unreasonable to expect that a similar amount of time (which was equated here) is required for the two types of instruction. More time per session might have allowed for a pace that provided students more time to process and practice the individual components of the TB+EF condition, whereas the components of the TB condition were relatively easy to establish. Similarly, a longer duration of training might have allowed students to better internalize the principles of the TB+EF condition.

Measurement—Finally, there are significant challenges of measuring the EF/SRL components that may be responsible for change. That is, a stronger link between the training and the components responsible for a change in learning might be generated more clearly if the mechanisms underlying that change are more appropriately measured. At a conceptual level, the primary difference when designing the TB+EF condition was to draw students' attention explicitly to the need to be engaged with the text and to be prepared to understand the material. The TB+EF condition also sought to make reading an active process for students and to help them understand that reading a passage and understanding it are not identical. Future studies might both identify better ways to accomplish this same goal and more precisely measure the processes that the intervention targets, like preparation, text engagement, and reflection, as well as similar but more general process (such as attention). It is also possible that stronger TB+EF results would have been obtained if outcomes had included direct measures of the processes taught in this condition (e.g., What is an example of a reading goal? How might you “get ready” to read? How might you “check in” while reading?).

Beyond the above three key factors, other aspects of the training conditions may have affected results. It is possible that some unique and shared components of the two researcher-implemented conditions were less helpful than others. For example, neither summarization nor rereading may have been an optimal use of the time in the form they were delivered here, although rereading was used less as a reading strategy to improve comprehension per se and more to provide for students an uninterrupted time to digest the information prior to quizzing. Related to quizzing, perhaps a more active testing program could have improved the productivity of the intervention time (e.g., longer quizzes, working through the quizzes with the text, intermixing content from across text). However, the goal of the present work was to emphasize goal setting, evaluation/monitoring, and focused effort on text, rather than to elaborate parameters under which the testing effect is amplified.

A final note is that this study did not emphasize the training of specific EFs per se. EF is often construed as a relatively abstract, “cool” (emotionally neutral) cognitive skill. There has been recent interest in “hot” (emotionally valenced) EF, although this distinction is much more investigated at the preschool level (e.g., Welsh & Peterson, 2014). In that literature and despite mixed findings, it does appear that cool, rather than hot, functions are more related to academic performance (e.g., Kim, Nordling, Yoon, Boldt, & Kochanska, 2013; Willoughby, Kupersmidt, Voegler-Lee, & Bryant, 2011). Of the EFs as they are traditionally construed, working memory has clearly received the most attention in terms of skills that may be malleable. The present work specifically did not seek to train working memory, as the goal was to contextualize EF contributions within existing reading comprehension approaches for struggling readers and because there is limited support for the generalizability of working memory training studies for reading comprehension (e.g., Melby-Lervag & Hulme, 2013; Shipstead et al., 2012). Whether specific content-based interventions might build in more experimentally manipulated components may be a topic for future investigation (e.g., comparing an existing academic intervention with validity data for struggling students against one that modifies it to systematically decrease demands on working memory). Also, direct comparison or combinatorial training studies that utilize both validated academic interventions and cognitive training packages could also be informative, though more data here are needed (see Kearns & Fuchs, 2013, for a recent review). It may also be argued that decontextualized EF training should not be expected to generalize to functional outcomes such as academics and that the challenge is to structurally alter task demands so that whatever EF weaknesses pertain interfere less with those functional outcomes. Finally, using reading content to practice metacognitive thinking, or to build self-efficacy or motivation in this area, might serve to increase these “executive skills” even if things like working memory capacity are not increased. As a somewhat related example, one recent study found that reading intervention might exert an effect on subsequent behavioral attention (Roberts et al., in press).

Study Limitations

To some extent, limitations are reflected in the issues noted above; that is, the two active conditions could have been more strongly differentiated at the design stage, fewer but more expanded EF components could have been emphasized, and a closer alignment of mechanism and measurement could potentially have been achieved. The training study was also limited in size and thus could detect only moderate effect sizes. The actual training also was not lengthy (approximately 10 hr), though whether additional hours would yield stronger differential effects is an open question. Although the development of the primary reading outcome measures and quizzes was informed by the pilot procedure, more extensive development work could have promoted stronger psychometric properties of them specifically for struggling readers, which in turn could lead to greater sensitivity to change. The present study also did not include standardized reading outcomes, primarily because of the short duration of training. In future studies, combining both experimenter-developed and standardized measures could prove insightful. Measures also were not equated with one another. For the quizzes in particular, because the content of each was confounded with time, this lack of equating made examining performance over time difficult, particularly with missing data due to absences. Despite these weaknesses and the lack of strong findings, strengths of this study include (a) the alignment of conceptual models of EF and SRL in the design of the training; (b) the focus on manipulations designed to address each stage of a combined EF/SRL process; (c) the contrast against a strong, active treatment; and (d) the direct experimental test of how EF-related functions might potentially affect academic content.

Relations of reading, EF, and SRL—The present study showed variable, and generally weak, relationships among these measures. There is a relatively well-established relationship of working memory to reading comprehension (e.g., Cain et al., 2004; St. Clair-Thompson & Gathercole, 2006). Beyond that, there are surprisingly few prior data with which to compare the present results, particularly in the age range examined here. Where literature does exist, findings are mixed and not often large. For example, Booth, Boyle, and Kelly (2010) found a mean meta-analytic effect size of 0.57 between learning disabled and non–learning disabled groups on EF measures (which translates to a correlation of $r \sim .28$). An exception that showed more robust relations was that of Sesma et al. (2009) and Cutting et al. (2009) in their mixed sample of children for an aspect of planning. That same measure of planning (excess moves) did have more robust relationships with outcomes relative to the more standard measure of correct responses in this sample as well.

Relationships of SRL specifically to reading performance or academic performance are not widely known. Baird et al. (2009) found correlations that ranged from $r = .08$ to .16 between such measures and learning disability status (which were significant in their very large sample), while Braten et al. (2013) found correlations with science text comprehension of $r = .28$ and .23 for background knowledge and science self-efficacy, respectively. Thus, the range of relationships in the present study are consistent with this prior work (see also Diseth, 2011). A recent study (Wolters et al., 2014) showed that in 7th to 12th graders, the largest relations between reading comprehension were evidenced for self-efficacy, perceived difficulty, perceived control, and anxiety, with correlations ranging from $r = |.27$ to $.37|$.

The variability of the relationship of EF and SRL measures with reading here is unlikely to be due to these measures' poorly representing the constructs of interest, for at least three reasons. First, it was not that relationships did not obtain—they did, and they were directionally appropriate—it was that their strength was less than expected and less consistent. Second, whatever relationships were found are consistent with extant literature. Third, the nature of the sample may have affected results (see Tables 3 and 6). For example, this study included only struggling readers (standardized reading performances ranged from M = 73.71 to 83.72 across four measures, with Letter-Word Identification M = 91.96), and students also had lower performances on measures such as D-KEFS Verbal Fluency (Mean SS = 7.86). As a result, the restricted range and variability (relative to a more heterogeneous sample) may have limited the size of correlations; such attenuation is well known under these conditions (Crocker & Algina, 1986; Goodwin & Leech, 2006).

The most consistent relations of EF or self-regulation to academic skill in the literature come from scales whose items overlap with the symptomatology of attention-deficit hyperactivity disorder, although stronger relations with academic functions are evident with regard to inattention rather than to hyperactivity/impulsivity (e.g., Cain & Bignell, 2014; Tymms & Merrell, 2011). Such relationships are interesting in terms of (a) the limited relationship between performance and rating scale measures (e.g., Toplak, West, & Stanovich, 2013), (b) the fact that cognitive and behavioral inattention are rarely examined together, and (c) the previously mentioned finding regarding reading effects on attention (Roberts et al., in press). Assessing constructs from both perspectives may shed light on how they individually and collectively influence academic performance.

While we did not expect powerful relationships among EF and SRL measures, we did assess several key aspects of the process of EF/SRL. For example, we included measures of goal setting (e.g., Titles, HWD, Tower of London), generation (e.g., Verbal Fluency, Titles), and monitoring/reflection (e.g., HDD, Effort/Efficacy). Despite this, we found that EF and SRL measures were relatively weakly interrelated, and the larger correlations appeared more sporadic than following a clear pattern. There are few studies that evaluate the relationship of neuropsychological measures of EF to self-regulation measures beyond preschool, though where examined, such relationships are generally small (e.g., Duckworth & Kern, 2011). Additional examination of the constructs of EF and SRL will clarify the nature of the relationships between these constructs. Within EF, and within SRL, relationships also were not strong. Particularly for EF, in addition to the weak relations between cognitive and performance measures as noted above, there are also generally weak relationships among performance components of EF especially at the preschool level (e.g., Willoughby, Pek, & Blair, 2013). Clearly, further work in elucidating these relationships is needed.

The results of this study focus attention on identifying both the significance and the magnitude of the relationships in determining what individual differences are important for reading comprehension. There remains a strong need for studies that address the structure of EF, particularly in the middle to late elementary range, to elaborate its relationships to both conceptually similar constructs (e.g., self-regulation) and academic skill, to do so at a broader (e.g., latent, meta-analytic) level, and to do so in the context of additional skills (Willoughby, Kupersmidt, & Voegler-Lee, 2012).

Conclusions

The present study showed improvements in student learning in researcher-implemented training conditions associated with taught content, though with very few differential effects related to the addition of EF/SRL components. There were variable but overall weak relationships among measures of reading, EF, and SRL. This study represents an initial attempt to use a variety of literatures to help make practical differences in students' learning. Future studies might utilize the more promising elements here to enhance, or at least better parameterize, the potential that EFs, broadly construed, have for improving reading comprehension.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Award Number P50 HD052117, Texas Center for Learning Disabilities, from the Eunice Kennedy Shriver National Institute of Child Health & Human Development to the University of Houston. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health & Human Development or the National Institutes of Health.

References

Antoniou F, Souvignier E. Strategy instruction in reading comprehension: An intervention study for students with learning disabilities. Learning Disabilities: A Contemporary Journal. 2007; 5(1):41–57.

Baird GL, Scott WD, Dearing E, Hamill SK. Cognitive self-regulation in youth with and without learning disabilities: Academic self-efficacy, theories of intelligence, learning vs. performance goal preferences, and effort attributions. Journal of Social and Clinical Psychology. 2009; 28(7):881–908.

Berkeley S, Mastropieri MA, Scruggs TE. Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. Journal of Learning Disabilities. 2011; 44(1):18–32. [PubMed: 21335506]

Berkeley S, Scruggs TE, Mastropieri MA. Reading comprehension instruction for students with learning disabilities, 1995–2006: A meta-analysis. Remedial and Special Education. 2010; 31(6): 423–436.

Best JR, Miller PH, Naglieri JA. Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. Learning and Individual Differences. 2011; 21(4):327–336. DOI: 10.1016/j.lindif.2011.01.007 [PubMed: 21845021]

Booth JN, Boyle JM, Kelly SW. Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. British Journal of Developmental Psychology. 2010; 28(1):133–176. [PubMed: 20306629]

Braten I, Ferguson LE, Anmarkrud O, Stromso HI. Prediction of learning and comprehension when adolescents read multiple texts: The roles of word-level processing, strategic approach, and reading motivation. Reading and Writing. 2013; 26(3):321–348.

Cain K, Bignell S. Reading and listening comprehension and their relation to inattention and hyperactivity. British Journal of Educational Psychology. 2014; 84(1):108–124. [PubMed: 24547756]

Cain K, Oakhill JV. Inference making ability and its relation to comprehension failure in young children. Reading and Writing. 1999; 11(5/6):489–503.

Cain K, Oakhill J, Bryant P. Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. Journal of Educational Psychology. 2004; 96(1):31.

Caprara GV, Fida R, Vecchione M, Del Bove G, Vecchio GM, Barbaranelli C, Bandura A. Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. Journal of Educational Psychology. 2008; 100(3):525.

Carretti B, Borella E, Cornoldi C, De Beni R. Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. Learning and Individual Differences. 2009; 19(2):245–251. DOI: 10.1016/j.lindif.2008.10.002

Cartwright KB. Insights from cognitive neuroscience: The importance of executive function for early reading development and education. Early Education & Development. 2012; 23(1):24–36.

Catts, HW., Hogan, TP., Adlof, SM. Developmental changes in reading and reading disabilities. In: Catts, HW., Kamhi, AG., editors. The connections between language and reading disabilities. Mahwah, NJ: Laurence Erlbaum; 2005. p. 25-40.

Chard DJ, Vaughn S, Tyler BJ. A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. Journal of Learning Disabilities. 2002; 35(5):386–406. [PubMed: 15490537]

Christopher ME, Miyake A, Keenan JM, Pennington B, DeFries JC, Wadsworth SJ, Olson RK. Predicting word reading and comprehension with executive function and speed measures across development: A latent variable analysis. Journal of Experimental Psychology: General. 2012; 141(3):470–488. DOI: 10.1037/a0027375 [PubMed: 22352396]

Cohen MJ, Morgan AM, Vaughn M, Riccio CA, Hall J. Verbal fluency in children: Developmental issues and differential validity in distinguishing children with attention-deficit hyperactivity disorder and two subtypes of dyslexia. Archives of Clinical Neuropsychology. 1999; 14(5):433–443. [PubMed: 14590585]

Compton DL, Fuchs D, Fuchs LS, Elleman AM, Gilbert JK. Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. Learning and Individual Differences. 2008; 18(3):329–337.

Crocker, L., Algina, J. Introduction to classical and modern test theory. New York, NY: Holt, Rinehart & Winston; 1986.

Cromley JG, Azevedo R. Testing and refining the direct and inferential mediation model of reading comprehension. Journal of Educational Psychology. 2007; 99(2):311.

Cutting LE, Materek A, Cole CA, Levine TM, Mahone EM. Effects of fluency, oral language, and executive function on reading comprehension performance. Annals of Dyslexia. 2009; 59(1):34–54. [PubMed: 19396550]

Delis, D., Kaplan, E., Kramer, J. Delis-Kaplan executive function scale. San Antonio, TX: Psychological Corp; 2001.

Dignath C, Buettner G, Langfeldt HP. How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. Educational Research Review. 2008; 3(2):101–129.

Dimitrov DM, Rumrill RD. Pretest-posttest designs and measurement of change. Work. 2003; 20:159–165. [PubMed: 12671209]

Diseth A. Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. Learning and Individual Differences. 2011; 21(2):191–195.

Duckworth AL, Kern ML. A meta-analysis of the convergent validity of self-control measures. Journal of Research in Personality. 2011; 45(3):259–268. [PubMed: 21643479]

Edmonds MS, Vaughn S, Wexler J, Reutebuch C, Cable A, Tackett KK, Schnakenberg JW. A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. Review of Educational Research. 2009; 79(1):262–300. [PubMed: 20072704]

Elbro C, Buch-Iversen I. Activation of background knowledge for inference making: Effects on reading comprehension. Scientific Studies of Reading. 2013; 17(6):435–452.

English L, Barnes MA, Fletcher JM, Dennis M, Raghubar KP. Effects of reading goals on reading comprehension, reading rate, and allocation of working memory in children and adolescents with spina bifi da meningomyelocele. Journal of the International Neuropsychological Society. 2010; 16:517–525. [PubMed: 20338082]

Fuchs LS, Fuchs D, Powell SR, Seethaler PM, Cirino PT, Fletcher JM. Intensive intervention for students with mathematics disabilities: Seven principles for effective practice. Learning Disability Quarterly. 2008; 31:79–92. [PubMed: 18815627]

Goodwin LD, Leech NL. Understanding correlation: Factors that affect the size of *r*. Journal of Experimental Education. 2006; 74(3):249–266.

Gough PB, Tunmer WE. Decoding, reading, and reading disability. Remedial and Special Education. 1986; 7(1):6–10.

Guthrie JT, Wigfield A, Metsala JL, Cox KE. Motivational and cognitive predictors of text comprehension and reading amount. Scientific Studies of Reading. 1999; 3(3):231–256.

Hornstra L, van der Veen I, Peetsma T, Volman M. Developments in motivation and achievement during primary school: A longitudinal study on group-specific differences. Learning and Individual Differences. 2013; 23:195–204.

Ilkowska, M., Engle, RW. Working memory capacity and self-regulation. In: Hoyle, RH., editor. Handbook of personality and self-regulation. Chichester, UK: Wiley; 2010. p. 263-290.

Johnson L, Graham S, Harris KR. The effects of goal setting and self-instruction on learning a reading comprehension strategy: A study of students with learning disabilities. Journal of Learning Disabilities. 1997; 30(1):80–91. [PubMed: 9009873]

Jurado MB, Rosselli M. The elusive nature of executive functions: A review of our current understanding. Neuropsychology Review. 2007; 17(3):213–233. [PubMed: 17786559]

Kamil, ML., Borman, GD., Dole, J., Kral, CC., Salinger, T., Torgesen, J. Improving adolescent literacy: Effective classroom and intervention practices. Jessup, MD: National Center for Education Evaluation and Regional Assistance; 2008. IES Practice Guide, NCEE 2008-4027

Karpicke JD, Roediger HL. The critical importance of retrieval for learning. Science. 2006; 319:966–968.

Kearns DM, Fuchs D. Does cognitively focused instruction improve the academic performance of low-achieving students? Exceptional Children. 2013; 79(3):263–290.

Kim S, Nordling JK, Yoon JE, Boldt LJ, Kochanska G. Effortful control in “hot” and “cool” tasks differentially predicts children’s behavior problems and academic performance. Journal of Abnormal Child Psychology. 2013; 41(1):43–56. [PubMed: 22798038]

Kuhn, MR., Stahl, SA. Fluency: A review of developmental and remedial practices. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement; 2000. Report No. 2-008

Lezak, MD., Howieson, DB., Bigler, ED., Tranel, D. Neuropsychological assessment. 5th. New York, NY: Oxford University Press; 2012.

Locascio G, Mahone EM, Eason SH, Cutting LE. Executive dysfunction among children with reading comprehension deficits. Journal of Learning Disabilities. 2010; 43(5):441–454. DOI: 10.1177/002221940935547 [PubMed: 20375294]

Logan S, Medford E, Hughes N. The importance of intrinsic motivation for high and low ability readers’ reading comprehension performance. Learning and Individual Differences. 2011; 21(1):124–128.

MacGinitie, W., MacGinitie, R., Maria, K., Dreyer, L., Hughes, K. Gates-MacGinitie reading tests—4. Itasca, IL: Riverside; 2000.

Mason LH. Teaching students who struggle with learning to think before, while, and after reading: Effects of self-regulated strategy development instruction. Reading & Writing Quarterly. 2013; 29(2):124–144.

McGee A, Johnson H. The effect of inference training on skilled and less skilled comprehenders. Educational Psychology. 2003; 23(1):49–59.

Melby-Lervag M, Hulme C. Is working memory training effective? A meta-analytic review. Developmental Psychology. 2013; 49(2):270. [PubMed: 22612437]

Miranda A, Villaescusa MI, Vidal-Abarca E. Is attribution retraining necessary? Use of self-regulation procedures for enhancing the reading comprehension strategies of children with learning disabilities. Journal of Learning Disabilities. 1997; 30(5):503–512. [PubMed: 9293232]

National Reading Panel, U.S. National Institute of Child Health & Human Development. Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction—Reports of the subgroups. Washington, DC: National Institute of Child Health and Human Development, National Institutes of Health; 2000.

Nelson JM, Manset-Williamson G. The impact of explicit, self-regulatory reading comprehension strategy instruction on the reading-specific self-efficacy, attributions, and affect of students with reading disabilities. Learning Disability Quarterly. 2006; 29(3):213–230.

Pearson PD, Hansen J, Gordon C. The effect of background knowledge on young children's comprehension of explicit and implicit information. Journal of Literacy Research. 1979; 11(3): 201–209.

Pintrich P. The role of motivation in promoting and sustaining self-regulated learning. International Journal of Educational Research. 1999; 31:459–470.

Pintrich, P. The role of goal orientation in self-regulated learning. In: Boekaerts, M.Pintrich, PR., Zeidner, M., editors. Handbook of self-regulation. San Diego, CA: Academic Press; 2000. p. 451-502.

Rawson KA, Dunlosky J. Relearning attenuates the benefits and costs of spacing. Journal of Experimental Psychology: General. 2013; 142(4):1113. [PubMed: 23088488]

Roberts G, Rane S, Fall AM, Denton CA, Fletcher JM, Vaughn S. The impact of intensive reading intervention on level of attention in middle school students. Journal of Clinical Child & Adolescent Psychology. 2015; 44:942–953. [PubMed: 24885289]

Roediger, HL., Putnam, AL., Smith, MA. Ten benefits of testing and their applications to educational practice. In: Mestre, JP., Ross, BH., editors. The psychology of learning and motivation. Vol. 55. San Diego, CA: Elsevier; 2011. p. 1-36.

Savage R, Cornish K, Manly T, Hollis C. Cognitive processes in children's reading and attention: The role of working memory, divided attention, and response inhibition. British Journal of Psychology. 2006; 97(3):365–385. DOI: 10.1348/000712605X81370 [PubMed: 16848949]

Scammacca N, Fall A, Roberts G. Benchmarks for expected annual growth for students in the bottom quartile of the normative distribution. Journal of Research on Educational Effectiveness. 2015; 8:366–379. [PubMed: 26726300]

Scammacca NK, Roberts G, Vaughn S, Stuebing KK. A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. Journal of Learning Disabilities. 2013; 48:369–390. doi: 0022219413504995. [PubMed: 24092916]

Schunk DH. Goal setting and self-efficacy during self-regulated learning. Educational Psychologist. 1990; 25(1):71–86.

Schunk DH. Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. Reading & Writing Quarterly. 2003; 19(2):159–172.

Sesma HW, Mahone EM, Levine T, Eason SH, Cutting LE. The contribution of executive skills to reading comprehension. Child Neuropsychology. 2009; 15(3):232–246. DOI: 10.1080/09297040802220029 [PubMed: 18629674]

Shallice T. Specific impairments of planning. Philosophical Transactions of the Royal Society of London. B, Biological Sciences. 1982; 298(1089):199–209. DOI: 10.1098/rstb.1982.0082 [PubMed: 6125971]

Shanahan, T., Callison, K., Carriere, C., Duke, NK., Pearson, PD., Schatschneider, C., Torgesen, J. Improving reading comprehension in kindergarten through 3rd grade: IES Practice Guide. Washington, DC: What Works Clearinghouse; 2010. NCEE 2010-4038

Shell DF, Murphy CC, Bruning RH. Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. Journal of Educational Psychology. 1989; 81(1):91.

Shipstead Z, Redick TS, Engle RW. Is working memory training effective? Psychological Bulletin. 2012; 138(4):628. [PubMed: 22409508]

Sideridis GD, Morgan PL, Botsas G, Padeliadu S, Fuchs D. Predicting LD on the basis of motivation, metacognition, and psychopathology: An ROC analysis. Journal of Learning Disabilities. 2006; 39(3):215–229. [PubMed: 16724794]

Sikora DM, Haley P, Edwards J, Butler RW. Tower of London Test performance in children with poor arithmetic skills. Developmental Neuropsychology. 2002; 21(3):243–254. DOI: 10.1207/S15326942DN2103_2 [PubMed: 12233937]

Solheim OJ. The impact of reading self-efficacy and task value on reading comprehension scores in different item formats. Reading Psychology. 2011; 32(1):1–27.

Spörer N, Schünemann N. Improvements of self-regulation procedures for fifth graders' reading competence: Analyzing effects on reading comprehension, reading strategy performance, and motivation for reading. Learning and Instruction. 2014; 33:147–157.

St Clair-Thompson HL, Gathercole SE. Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. Quarterly Journal of Experimental Psychology. 2006; 59(4):745–759. DOI: 10.1080/17470210500162854

Stern SK, Morris MK. Discrimination of ADHD and reading disability in adults using the D-KEFS. Archives of Clinical Neuropsychology. 2013; 28:125–134. [PubMed: 23246933]

Stroud, KC. Unpublished doctoral dissertation. Texas A&M University; College Station: 2006. Development of the school motivation and learning strategies inventory.

Swanson JM, Schuck S, Porter MM, Carlson C, Hartman CA, Sergeant J, Wigal T. Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and SWAN rating scales. International Journal of Educational Psychological Assessment. 2012; 10(1):51–70. [PubMed: 26504617]

Swanson JM, Sergeant JA, Taylor E, Sonuga-Barke EJS, Jensen PS, Cantwell DP. Attention-deficit hyperactivity disorder and hyperkinetic disorder. Lancet. 1998; 351(9100):429–433. [PubMed: 9482319]

Taboada A, Tonks SM, Wigfield A, Guthrie JT. Effects of motivational and cognitive variables on reading comprehension. Reading and Writing. 2009; 22(1):85–106.

Therrien WJ. Fluency and comprehension gains as a result of repeated reading a meta-analysis. Remedial and Special Education. 2004; 25(4):252–261.

Toplak ME, West RF, Stanovich KE. Do performance-based measures and ratings of executive function assess the same construct? Journal of Child Psychology and Psychiatry. 2013; 54(2):131–143. DOI: 10.1111/jcpp.12001 [PubMed: 23057693]

Torgesen JK. Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. Learning Disabilities Research & Practice. 2000; 15(1):55–64.

Torgesen, JK., Wagner, R., Rashotte, C. Test of word reading efficiency. 2nd. Austin, TX: PRO-ED; 1999.

Tymms P, Merrell C. ADHD and academic attainment: Is there an advantage in impulsivity? Learning and Individual Differences. 2011; 21(6):753–758.

Wagner, RK., Torgesen, JK., Rashotte, CA., Pearson, NA. Test of sentence reading efficiency and comprehension (TOSREC). Austin, TX: PRO-ED; 2009.

Wanzek J, Vaughn S, Scammacca NK, Metz K, Murray CS, Roberts G, Danielson L. Extensive reading interventions for students with reading difficulties after Grade 3. Review of Educational Research. 2013; 83(2):163–195. doi: 0034654313477212.

Welsh M, Peterson E. Issues in the conceptualization and assessment of hot executive functions in childhood. Journal of the International Neuropsychological Society. 2014; 20(2):152–156. [PubMed: 24468077]

Willoughby M, Kupersmidt J, Voegler-Lee M, Bryant D. Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. Developmental Neuropsychology. 2011; 36(2):162–180. [PubMed: 21347919]

Willoughby MT, Kupersmidt JB, Voegler-Lee ME. Is preschool executive function causally related to academic achievement? Child Neuropsychology. 2012; 18(1):79–91. [PubMed: 21707258]

Willoughby MT, Pek J, Blair CB. Measuring executive function in early childhood: A focus on maximal reliability and the derivation of short forms. Psychological Assessment. 2013; 25(2):664. [PubMed: 23397928]

Wolters CA, Denton CA, York MJ, Francis DJ. Adolescents' motivation for reading: Group differences and relation to standardized achievement. Reading and Writing. 2014; 27(3):503–533.

Woodcock, RW., McGrew, KS., Mather, N. The Woodcock-Johnson III tests of achievement. Atasca, IL: Riverside; 2001.

Ylvisaker M, Feeney T. Executive functions, self-regulation, and learned optimism in paediatric rehabilitation: A review and implications for intervention. Developmental Neurorehabilitation. 2002; 5(2):51–70.

Zelazo PD, Carter A, Reznick J, Frye D. Early development of executive function: A problem solving framework. Review of General Psychology. 1997; 1(2):198–226.

Zentall SS, Lee J. A reading motivation intervention with differential outcomes for students at risk for reading disabilities, ADHD, and typical comparisons: "Clever is and clever does. Learning Disability Quarterly. 2012; 35:248–259. doi: 0731948712438556.

Zhou Q, Chen SH, Main A. Commonalities and differences in the research on children's effortful control and executive function: A call for an integrated model of self-regulation. Child Development Perspectives. 2012; 6(2):112–121.

Zimmerman BJ. Self-efficacy: An essential motive to learn. Contemporary Educational Psychology. 2000; 25(1):82–91. [PubMed: 10620383]

Zimmerman, BJ., Schunk, DH. Self-regulated learning: An introduction and an overview. In: Zimmerman, BJ., Schunk, DH., editors. Handbook of self-regulated learning and performance. New York, NY: Routledge; 2011. p. 65-84.

Zuffiano A, Alessandri G, Gerbino M, Luengo Kanacri BP, Di Giunta L, Milioni M, Caprara GV. Academic achievement: The unique contribution of self-efficacy beliefs in self-regulated learning beyond intelligence, personality traits, and self-esteem. Learning and Individual Differences. 2013; 23:158–162.

Table 1

Instructional Activities and Structure for Two Active Trainings in Pilot and Experiment.

Group	Activity content	Section	Time[a]	Description	Pilot	Experiment
TB	Word skim	Preread	1.5	Ss identify and Ts read unknown words prior to reading	Yes	Yes
	Text-based questions	Postread	8.0	Ts ask Ss questions about events in text and scaffold responses	Yes	Yes
TB+EF	Titles	Preread	2.5	Ss discuss what they know about topic based on title	Yes	Yes
	Goals	Preread	1.5	Ss choose a goal stem for the group; Ss plan to find what the passage is “mainly about”	Yes	Yes
	Get ready	Preread	2.0	Ts remind Ss to prepare to read, register goal, and to check-in during reading	Yes	Yes
	Check-in	During	2.0	Ss indicate whether they “checked-in” during reading	Yes (written checkmarks)	Yes (via discussion)
	Goal-check	Postread	2.0	Ts ask Ss if they met group goal and why; Ts/Ss discuss	Yes	No
Both conditions	Vocabulary cards	Preread	5.0 and 2.0	Ts introduce definition, related words, examples, plus discussion	Yes	Yes
	Potentially unknown words	Preread	2.0	Ss select words to be defined; Ts provide definition	Yes (TB+EF only)	Yes
	Oral reading	During	5.0	Ss take turns reading aloud, with assistance from Ts as needed	yes	yes
	Summarize	Postread	6.0 and 8.0	Ss summarize passage aloud with scaffolding from Ts	yes	yes
	Silent rereading	Postread	5.0	Ss reread text silently	Yes (TB only)	Yes
	Quizzing	Postread	5.0	Ss answer multiple-choice question about the text	Yes (TB+EF only)	Yes
	Motivation	Throughout	Not applicable	Ts monitor behavior, offer reward opportunities	Yes	Yes

Note. TB = text-based reading; S = student; T = tutor; TB+EF = text-based reading plus executive function.

[a]Approximate time in minutes (though not strictly timed); for cells with two values, minutes for TB are listed first, and minutes for TB+EF are listed second.

Table 2

Demographic Data for Experiment Study By Group.

	Total				CON				TB				TB+EF			
Demographic	***n***	***M***	***SD***	**%**	***n***	***M***	***SD***	**%**	***n***	***M***	***SD***	**%**	***n***	***M***	***SD***	**%**
Age	75	10.18	0.58		27	10.02	0.47		24	10.22	0.56		24	10.32	0.68	
Sex (% female)				37.33				51.85				37.50				20.83
Ethnicity (%)																
African American				40.00				44.44				33.33				41.67
Hispanic				54.67				48.15				67.77				50.00
Other				5.33				7.41				NA				8.33

Note. CON = control; TB = text-based reading; TB+EF = text-based reading plus executive function. NA = not applicable.

Table 3

Descriptive Statistics for the Pretest Measures for Experiment.

	Total			CON			TB			TB+EF		
Measure	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
GMPC	72	81.17	6.45	27	82.56	5.71	21	80.33	6.84	24	80.33	6.88
WJ-III LWID	74	91.96	10.69	26	90.85	9.55	24	92.67	11.05	24	92.46	11.79
TOWRE-2 SWE	74	83.73	11.34	26	86.38	12.53	24	83.63	11.20	24	80.96	9.79
WJ-III PC	74	83.54	7.05	26	83.27	7.50	24	83.38	7.12	24	84.00	6.77
TOSREC	75	73.71	10.69	27	74.74	12.63	24	72.40	10.46	24	73.88	8.66

Note. CON = control; TB = text-based reading; TB+EF = text-based reading plus executive function; GMPC = *Gates-MacGinitie Reading Tests*–Comprehension; WJ-III LWID = *Woodcock-Johnson III Test of Achievement* (3rd ed.) Letter Word Identification; TOWRE-2 SWE = *Test of Word Reading Efficiency* (2nd ed.), Sight Word Efficiency subtest; WJ-III PC = *Woodcock-Johnson III Test of Achievement* (3rd ed.) Passage Comprehension; TOSREC = *Test of Sentence Reading Efficiency and Comprehension*. All scores are standard scores.

Table 4

Descriptive Statistics of Measures Given at Pre- and Posttest for Experiment.

	Total			CON			TB			TB+EF		
Statistic	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
Age (years)	75	10.18	0.58	27	10.02	0.47	24	10.22	0.56	24	10.32	0.68
Attendance (days)	NA	NA	NA	NA	NA	NA	24	16.71	2.77	24	16.75	3.19
Lessons	NA	NA	NA	NA	NA	NA	24	15	1.14	24	14.73	1.20
Minutes	NA	NA	NA	NA	NA	NA	24	533.96	79.83	24	550.13	86.35
Pretest proximal	74	0.38	0.18	26	0.41	0.20	24	0.36	0.16	24	0.36	0.18
Posttest proximal	75	0.45	0.19	27	0.38	0.19	24	0.52	0.17	24	0.47	0.20
Posttest related	75	0.43	0.18	27	0.39	0.16	24	0.47	0.21	24	0.43	0.16
Posttest unrelated	75	0.49	0.24	27	0.49	0.23	24	0.47	0.23	24	0.50	0.27
Background (pre)	74	2.62	1.56	26	2.88	1.93	24	2.29	1.20	24	2.67	1.43
Background (post)	75	4.49	1.95	27	3.30	1.75	24	5.08	1.84	24	5.25	1.67

Note. CON = control; TB = text-based reading; TB+EF = text-based reading plus executive function; Attendance = in days (maximum 20); Lessons = number covered by active conditions (maximum 16); Minutes = total duration of training for active conditions. Pretests and posttests are in percentage correct. Background is raw score (maximum 10). NA = not applicable.

Table 5

Descriptive Statistics of Self-Regulated Learning Measures for All Participants for Experiment.

	Total			CON			TB			TB+EF		
Measure	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
Titles[a]												
Known state	63	13.32	4.36	20	13.50	4.93	21	11.90	4.10	22	14.50	3.84
Known reasonable	63	8.10	3.16	20	7.70	3.56	21	7.95	3.67	22	8.59	2.20
Potential state	63	10.40	5.27	20	10.85	5.71	21	9.05	4.22	22	11.27	5.73
Potential reasonable	63	8.46	4.88	20	8.45	5.38	21	7.81	3.86	22	9.09	5.38
HWD[b]												
Pretest	74	0.73	0.18	26	0.76	0.16	24	0.72	0.20	24	0.70	0.18
Posttest	74	0.76	0.15	27	0.81	0.11	24	0.74	0.15	23	0.72	0.17
HDD[c]												
Pretest	74	0.71	0.18	26	0.79	0.14	24	0.70	0.17	24	0.65	0.21
Posttest	74	0.76	0.14	27	0.78	0.12	24	0.76	0.16	23	0.73	0.14
SLQ[d]												
Effort/Efficacy (pretest)	74	28.65	6.81	26	30.46	5.24	24	28.17	8.35	24	27.17	6.44
Effort/Efficacy (posttest)	72	27.49	7.73	25	27.48	9.07	24	27.83	7.33	23	27.13	6.82
Skill/Preference (pretest)	74	8.41	3.31	26	8.50	3.25	24	9.25	2.77	24	7.46	3.72
Skill/Preference (posttest)	72	8.13	3.40	25	7.64	3.46	24	8.79	2.98	23	7.96	3.76
SWAN[e]												
Inattention	74	5.82	9.79	27	4.70	11.07	23	4.39	10.04	24	8.46	7.65
Hyper/Impulsive	74	4.15	8.67	27	4.52	9.10	23	1.65	4.91	24	6.13	10.55

Note. CON = control; TB = text-based reading; TB+EF = text-based reading plus executive function; HWD = How Will I Do?; HDD = How Did I Do?; SLQ = *Student Learning Questionnaire*; SWAN = *Strengths and Weaknesses of ADHD Symptoms and Normal Behavior*. SWAN and Titles were administered only at posttest.

[a]See text for description (raw scores are presented; *stated* refers to number of ideas student identified and *reasonable* to the number of those ideas that are in fact related to content).

[b]In proportion of 100%; see text for description.

[c]In proportion of 100%; see text for description.

[d]Effort/Efficacy and Skill/Preference are raw scores, maximum ranges are 0 to 39 and 0 to 15, respectively; see text for description.

[e]Range of scale is −27 to +27; see text for description.

Table 6

Descriptive Statistics of Executive Function Measures for Experiment.

	Total			CON			TB			TB+EF		
Measure	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
TOL (Total)	71	12.82	4.09	26	11.88	4.49	21	14.38	2.82	24	12.46	4.31
TOL (Excess Moves)	71	7.24	6.74	26	7.00	5.73	21	6.57	6.70	24	8.08	7.90
D-KEFS												
Letter Fluency (scale score)	71	7.86	2.79	26	7.65	2.59	21	7.76	3.03	24	8.17	2.88
Category Fluency (scale score)	71	8.56	2.60	26	8.77	2.53	21	8.05	2.42	24	8.79	2.86
Switching (scale score)	71	6.85	3.28	26	6.62	2.77	21	6.10	3.70	24	7.75	3.33

Note. CON = control; TB = text-based reading; TB+EF = text-based reading plus executive function; TOL = Tower Task; D-KEFS = *Delis-Kaplan Executive Function System*. TOL: total raw score, maximum 20; excess moves, range is 0 to 28.

Table 7

Correlations Among Executive Function (Pretest) and Self-Regulated Learning Measures (Pretest and Posttest).

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. TOL (total)	—													
2. TOL (excess moves)	−0.19	—												
3. D-KEFS Letter Fluency (SS)	−0.04	−0.14	—											
4. D-KEFS Category Fluency (SS)	0.07	< 0.01	**0.56****	—										
5. D-KEFS Verbal Switch (SS)	0.06	< 0.01	**0.39****	**0.29***	—									
6. HWD (pretest)	−0.18	−0.03	−0.06	−0.13	−0.21	—								
7. HDD (pretest)	−0.24*	−0.01	−0.06	−0.14	−0.28*	**0.76****	—							
8. SLQ Effort/Efficacy (pretest)	−0.07	0.13	0.14	0.05	<0.01	0.25*	0.33*	—						
9. SLQ Skill/Preference (pretest)	−0.09	0.19	< 0.01	−0.17	−0.33*	0.37*	0.38**	**0.29***	—					
10. Titles (known state)	−0.32*	0.23	0.05	−0.02	−0.03	0.06	0.09	0.04	0.08	—				
11. Titles (known reasonable)	−0.14	0.05	−0.04	−0.10	−0.13	0.07	0.14	−0.06	0.14	**0.62****	—			
12. Titles (potential state)	−0.17	0.01	0.10	−0.15	−0.12	0.20	0.28*	0.18	0.29*	**0.63****	**0.55****	—		
13. Titles (potential reasonable)	−0.13	−0.10	0.02	−0.22	−0.11	0.10	0.18	0.12	0.20	**0.38***	**0.50****	**0.84****	—	
14. SWAN Inattention	0.09	0.23	−0.01	0.05	0.25*	0.10	−0.09	−0.33*	−0.03	0.04	−0.15	−0.11	−0.24	—
15. SWAN Hyperactive/Impulsive	0.11	0.20	0.21	0.28*	0.20	< 0.01	< 0.01	−0.08	−0.05	0.13	−0.08	−0.04	−0.16	**0.71****

Note. See notes for Tables 5 and 6 for explanation of abbreviations. TOL and D-KEFS administered only at pretest; Titles and SWAN administered only at posttest. SS = standard score. Boldface type indicates correlations of components of same test/within same battery.

* $p < .05$.

** $p < .001$.