UC Irvine
UC Irvine Previously Published Works

Title

Examining the Efficacy of Targeted Component Interventions on Language and Literacy for Third and Fourth Graders Who are at Risk of Comprehension Difficulties

Permalink

https://escholarship.org/uc/item/18v5j7cq

Journal

Scientific Studies of Reading, 22(6)

ISSN

1088-8438

Authors

Connor, Carol McDonald
Phillips, Beth M
Kim, Young–Suk Grace
et al.

Publication Date

2018-11-02

DOI

10.1080/10888438.2018.1481409

Peer reviewed

Examining the Efficacy of Targeted Component Interventions on Language and Literacy for Third and Fourth Graders Who are at Risk of Comprehension Difficulties

Carol McDonald Connor, Beth M. Phillips, Young–Suk Grace Kim, Christopher J. Lonigan, Michael P. Kaschak, Elizabeth Crowe, Jennifer Dombek & Stephanie Al Otaiba

Routledge
Taylor & Francis Group

Check for updates

Examining the Efficacy of Targeted Component Interventions on Language and Literacy for Third and Fourth Graders Who are at Risk of Comprehension Difficulties

Carol McDonald Connor[a], Beth M. Phillips[b], Young–Suk Grace Kim[a], Christopher J. Lonigan[b], Michael P. Kaschak[b], Elizabeth Crowe[c], Jennifer Dombek[b], and Stephanie Al Otaiba[d]

[a]University of California, Irvine; [b]Florida Center for Reading Research, Florida State University; [c]University Academy, Panama City, Florida; [d]Southern Methodist University

ABSTRACT

Testing a component model of reading comprehension in a randomized controlled trial, we evaluated the efficacy of 4 interventions that were designed to target components of language and metacognition that predict children's reading comprehension: vocabulary, listening comprehension, comprehension of literate language, academic knowledge, and comprehension monitoring. Third- and 4th-graders with language skills falling below age expectations participated ($N = 645$). Overall, the component interventions were only somewhat effective in improving the targeted skills, compared to a business-as-usual control (g ranged from −.14 to .33), and no main effects were significant after correcting for multiple comparisons. Effects did not generalize to other language skills or to students' reading comprehension. Moreover, there were Child Characteristic × Treatment interaction effects. For example, the intervention designed to build sensorimotor mental representations was more effective for children with weaker vocabulary skills. Implications for component models of reading and interventions for children at risk of reading comprehension difficulties are discussed.

Improving reading comprehension skills has proven to be more difficult than anticipated (e.g., James-Burdumy et al., 2009). In an effort to understand why, the Reading for Understanding Network was formed to take findings from basic research; translate this research into meaningful interventions to improve children's reading for understanding; and then evaluate, in randomized controlled trials, whether the new interventions were effective in improving reading comprehension. The aims are important because, according to the 2015 National Assessment of Educational Progress (https://nces.ed.gov/nationsreportcard/), only 36% of fourth graders in the United States are reading at or above proficient levels. Moreover, for minority children (e.g., African American, Hispanic) this rate is much lower: 18%–21%. It is also lower (21%) for children living in poverty (i.e., qualify for the National School Lunch Program [NSLP]). Of great concern, fully 44% of children eligible for the NSLP are reading below basic levels. The aims of the Florida Center for Reading Research Reading for Understanding Network were threefold: (a) to identify the underlying cognitive and linguistic processes that contribute to successful text comprehension and comprehension for the purpose of learning, (b) to rapidly develop appropriate interventions that enhance these processes and determine which are effective for students with varying skill constellations, and (c) to test the efficacy and

---

usability of these interventions. The randomized controlled study presented here was designed to meet these aims by evaluating the efficacy of interventions designed for third and fourth graders. The interventions targeted key linguistic and metacognitive skills. Basic research has revealed that these skills are associated with proficient reading comprehension (e.g., Hulme, Nash, Gooch, Lervåg, & Snowling, 2015).

Proficient reading comprehension is defined as the ability to "demonstrate a strong understanding of the text … to extend the ideas in the text by making inferences, drawing conclusions, and making connections to [students'] own experiences" (National Assessment Governing Board, 2007, p. 24). Reading comprehension requires children to fluently decode and understand what they are reading (Rand Study Group & Snow, 2001; Rapp, Van Den Broek, McMaster, Kendeou, & Espin, 2007; Scarborough, 2001). Basic processes underlying reading comprehension are complex and call on the oral language system and a conscious understanding of this system (i.e., metalinguistic and metacognitive awareness) at all levels from semantic and morpho-syntactic to pragmatic awareness (Morrison, Bachman, & Connor, 2005). Higher order metacognitive skills also predict comprehension (Efklides & Misailidi, 2010; Rapp et al., 2007). Thus accumulating evidence indicates that multiple skills, such as grasp of complex syntax, comprehension monitoring and comprehension strategy use (Baker, 2008), awareness of text structure (Williams, Stafford, Lauer, Hall, & Pollini, 2009), and background knowledge (Rapp et al., 2007), support creating coherent mental representations (or situation models; Kintsch, 1998; Van Den Broek, Risden, Fletcher, & Thorlow, 1996) and result in proficient reading for understanding.

The original framework for the design of the interventions reported here was the Florida Center for Reading Research Reading for Understanding component model of reading comprehension, which was developed to guide this network's projects. It builds on the simple view model (Hoover & Gough, 1990; Joshi & Aaron, 2000) and extends other component models (e.g., Mellard & Fall, 2012). In this model, reading for understanding requires fluent decoding and word-level skills, proficient vocabulary strong syntax/listening comprehension skills, and the ability to bring automatic skills to the conscious level when the reading task demands it—reflective comprehension processes (e.g., metacognition). All of these processes are supported by developing text-specific, linguistic, metalinguistic, and metacognitive processes (e.g., comprehension monitoring) in combination with the instruction children are receiving (Connor, 2016a; Kim, 2017). The theory of action is that if we can strengthen key component skills that contribute to proficient reading for understanding by explicitly teaching these skills through intervention, they can be learned and can become more fluent and automatic. Therefore, reading comprehension should also improve. In this study, the interventions focused on improving key component skills that are associated with stronger reading comprehension: text structure knowledge, oral language (including vocabulary, listening comprehension, complex syntax and narrative skills, and academic knowledge), and comprehension monitoring.

The intent of each of these interventions, part of Comprehension Tools for Teachers (CTT), was to provide effective Tier 2 instruction designed to improve one or more component skills that predict reading comprehension, within a Multi-tiered System of Support (Gersten & Dimino, 2006). In this system, also known as Response to Intervention, children whose reading skills do not improve as expected, even when receiving effective general education literacy instruction (i.e., Tier 1), are provided with more intense and targeted literacy instruction, often referred to as Tier 2. Tier 2 instruction is typically provided in smaller flexible learning groups of students with similar learning goals, multiple times per week, for about 20–30 min, which was the model used in these CTT interventions. Components considered in the CTT Component Interventions are described next.

Text structure knowledge

Text structure is the organization found in various types of texts (Graesser, Golding, & Long, 1991; Tompkins & McGee, 1989). For example, we know that the topic sentence in a well-crafted paragraph frames the content of the paragraph. Headings help us locate specific information.

There is accumulating evidence that when students are explicitly taught to attend to text structure, their comprehension skills improve. Text structure knowledge may aid in text comprehension because it provides schemas to organize and understand new information (Graesser et al., 1991; Stevens, Van Meter, & Warcholak, 2010; Tompkins & McGee, 1989; Williams et al., 2005; Williams et al., 2009). In narratives, important elements of a story (i.e., text structure) typically include main characters, the setting, an initiating event, a problem, attempts to solve the problem, and the resolution (Baker & Stein, 1981; Fitzgerald, 1989). Expository text structure (e.g., informational text), incorporates discourse devices such as "first, next, last" to identify sequences and "if, then" and "because" to identify logical and causal relations. Moreover, because fluent text structure knowledge may reduce cognitive load, one would predict that children's comprehension skills and knowledge gained from the text (i.e., academic knowledge) should be supported.

Oral language skills (vocabulary, syntax, semantic, and academic knowledge)

Children's oral language skills develop over time and increase in sophistication on a fairly predictable timetable, although there are individual child differences that are a function of biological and environmental influences (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Locke, 1997). Whereas humans may learn to read at any age, it is quite likely that the developing linguistic system will influence the rate and efficient development of children's literacy skills during the pre-K through elementary years (Scarborough, 1998). Reading and oral language are related throughout development (Catts, Fey, Zhang, & Tomblin, 1999; Hulme et al., 2015; Scarborough, 2001). Accumulating data reveal that their relations change as children mature (Kim & Wagner, 2015; Storch & Whitehurst, 2002) and are at times reciprocal (Connor et al., 2016), with experience in each potentially affecting skill acquisition in the other; however, this depends on children's age and the instruction they receive.

For third and fourth graders, foundational linguistic skills are generally well developed, although there is large variation among students; they are able to carry on relatively sophisticated conversations, monitor their understanding of the conversation (e.g., by asking questions), and have an emerging grasp of metalinguistics (e.g., idioms, puns; Caillies & Le Sourn-Bissaoui, 2008; Cain, Towse, & Knight, 2009). Evidence indicates that successful application of linguistic skills to the process of comprehending texts requires language and metalinguistic skills (Cain, Oakhill, & Bryant, 2004; Cromley & Azevedo, 2007; Kim, 2016, 2017; Kim & Pilcher, 2016). Competency at consciously applying these linguistic and metalinguistic skills to the task of reading comprehension continues to improve as children move from achieving fluent decoding skills to learning from text (Chall, 1996). Those children who are struggling to comprehend or who have weaker language skills may need more explicit instruction in complex language skills, including more advanced vocabulary, listening comprehension, comprehension of literate language, and academic/semantic knowledge (Clarke, Snowling, Trulove, & Hulme, 2010; van der Lely & Marshall, 2010).

Metacognitive skills, or the conscious thinking about thinking (Flavell, 1979), support students' reading for understanding. A key metacognitive skill explored in this study was comprehension monitoring, which is recognizing when one does not understand and then employing repair strategies to improve understanding (i.e., comprehension monitoring; Baker, 1984; Chen, 2009; Ruffman, 1999), thereby constructing coherent situation models and coherent mental representations of texts (e.g., Cain et al., 2004; Elliot-Faust & Pressley, 1986; Kim, 2015, 2016, 2017; Kim & Phillips, 2014; Rapp et al., 2007; Skarakis-Doyle, Dempsey, & Lee, 2008). When understanding breaks down, students should understand how to employ repair strategies. Another strategy considered here is creating sensorimotor simulations of text, which is associated with stronger reading comprehension for younger children (Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006). Consciously enacting simulations that include motor actions may support the grasp of more semantically and cognitively abstract ideas, such as the opposing forces that underlie arguments or interpersonal conflict (Lakoff & Johnson, 1980).

The component interventions

The component intervention included Comprehension Monitoring and Providing Awareness of Story Structure (COMPASS), Language in Motion, and Enacted Reading Comprehension (Enacted RC) in Grade 3, and Enacted RC and Teaching Expository Text Structure (TEXTS) in Grade 4.

COMPASS

COMPASS explicitly focuses on three component skills of reading comprehension: vocabulary, comprehension monitoring, and text structure knowledge (Marulis & Neuman, 2010). Each lesson has a scaffolded learning format of I-do, we-do, and you-do (Pearson & Gallagher, 1983); overall, the lessons are sequenced such that each component's instruction increased progressively in difficulty over time.

For comprehension monitoring, students heard a very short story (e.g., two sentences) with an accompanying illustration and were asked whether the story made sense. All the stories were told with illustrations in the first 4 weeks. Prompt illustrations were progressively removed in the second 4 weeks. At the end of each comprehension monitoring portion of the session, children were told that "when they listen to a story, things have to make sense to them, and if they do not, they have to stop and ask questions." The comprehension monitoring portion of the lesson took approximately 5 min (see Kim & Phillips, 2016, for a similar structure but for prekindergartners).

Vocabulary and text structure lessons were taught using researcher-developed narratives, which were written to have a clear text structure in terms of characters, setting, leading events, problem, and resolution. The stories, which focused on themes of "traditions in different cultures," included characters to whom participating children could relate in terms of age, gender, and racial backgrounds and aligned to third-grade social studies standards. Each story had four to six accompanying illustrations that showed the sequence of events. One story was used each week and accompanied by activities including reading aloud, dialogic reading questions, and retell. Key target words from every story were identified and explicitly taught in each lesson (e.g., tradition, miracle). Children were also explicitly taught about text structure—characters, setting, initiating events, problem, efforts to solve the problem, and resolution—using visual aids as well as a song. Retell was first modeled by interventionists, followed by opportunity for children to practice using story illustrations. Students read the text as they listened to the orally presented narrative; they were taught to mark text (e.g., circle) where they found target vocabulary words and information relevant to text structure. We anticipated that COMPASS would have an effect on comprehension monitoring, comprehension of literate language, listening comprehension, and vocabulary skills.

Language in motion

Language in Motion was designed to support student's *understanding and expressive use of literate language features* known to differentiate text from typical oral language (e.g., Benson, 2009; Zipoli, 2017), which can be difficult for children during both oral and reading comprehension (e.g., Megherbi & Ehrlich, 2005; Oakhill, Cain, & Bryant, 2003). Targeted language elements for third grade included relative clauses, passive sentence structures, and resolution of anaphoric pronoun references; one unit also focused on figurative language devices of metaphor, simile, and common idioms. In addition, each 4-week unit included lessons on targeted mental state vocabulary words, which were also embedded in the texts, focusing on the manner in which different words distinctly represented levels of certainty (i.e., speculate, conclude; e.g., Moore, Pure, & Furrow, 1990; Schwanenflugel, Fabricius, & Noyes, 1996). All Language in Motion lessons were embedded within a narrative adventure story (a haunted house in third grade) and science themes related to basic science concepts (e.g., in third grade, how sound and light travel in waves). Hands-on science activities and two-dimensional story characters and storyboards supported highly interactive, game-like lessons in which new language content was first modeled within oral narratives and activity instructions and then students were challenged to correctly respond behaviorally to these

instructions and to use new linguistic content in their verbal responses. Lesson scripts provided interventionists with examples of how to support scaffolding for students who found lessons difficult and upward extensions for students who more quickly mastered new content. We hypothesized that Language in Motion would have an effect on vocabulary, listening comprehension, and comprehension of literate language.

Enacted RC

In this intervention we applied the principles of embodied cognition, which hold that linguistic understanding emerges through the construction of sensorimotor simulations (Glenberg, Gutierrez, Levin, Japunitich, & Kaschak, 2004; Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006). For third and fourth graders, we conjectured that motor actions might support the grasp of more abstract and science-related ideas, such as the opposing forces that underlie earthquakes and weather, and in narrative, arguments or interpersonal conflict (Lakoff & Johnson, 1980), by helping students consciously use embodied cognition as a strategy for improving understanding of the texts (Connor et al., 2014; Glenberg, Brown, & Levin, 2007; Kaschak, Connor, & Dombek, 2017). In particular, we focused on systematically building students' understanding of abstract concepts starting with concrete simulation strategies to understand opposing forces in earthquakes (moving hands together to simulate the earthquakes; shaking a toy house full of toy furniture), then moving to more abstract concepts of opposing forces in hurricanes and tornados (waving arms/hands to simulate the circling winds), persuasive/argumentative texts (circling one hand for one side of the argument and circling the other hand for the opposing argument). For the final phase of the intervention, students read the novel *A Single Shard* (Park, 2001). These lessons focused on understanding internal conflicts and moral dilemmas faced by the main characters of the book in the context of opposable forces that could be illustrated with gestures. Lessons were presented in three phases. In Phase 1 (about three weeks), the most concrete phase, children read books on earthquakes, hurricanes, and tornados. In Phase 2 (about two weeks), children read and wrote persuasive texts focusing on the opposing forces of opinion. In Phase 3 (about five weeks) they read *A Single Shard* and discussed how moral dilemmas are like opposing forces and opposing opinions. Students each received a copy of *A Single Shard*, which they were allowed to keep. In this last phase, children read the book during lessons and at home. We hypothesized that Enacted RC would have an effect on children's academic knowledge through experience with science texts, vocabulary, and possibly their comprehension monitoring because sensori-motor simulations can be used as a comprehension repair strategy. However, comprehension monitoring was not a direct focus of the Enacted RC intervention.

Texts

The TEXTS intervention is designed to explicitly teach students about expository text structure and to emphasize how certain words can signal, or serve as a "clue," that an expository text has a particular structure. The predominant text structures that shaped TEXTS are sequencing, or the order of events; cause and effect, or why things occur; compare and contrast, or how things are the same or different; and problem and solution, or how to solve issues (Meyer & Poon, 2001).

Our intent was to have a series of units with different expository text structures to support struggling readers' comprehension of academic content. TEXTS books were written based on common topics and state standards for fourth grade (e.g., The Food Chain, The Scientific Method, Pioneer Homes, Making Good Choices). The design team developed a scope and sequence for each book unit, with scripted lesson plans for project interventionists to use with small groups. Each unit lasted 1 week. On the 1st day, interventionists spent the first 5 min introducing signal words and the notion of searching for certain words that provide a clue about the type of text structure (e.g., first, next, last for sequencing). Then students read the book for about 10 min, and the final 5–10 min involved application and practice. The "wrap-up" was brief, and interventionists asked students to review what they had learned (note this was the final stage for each day of instruction). On the 2nd day, interventionists and students played a signal, or clue word, memory game; guided students through another read of the story; and instructed

students how to use a graphic organizer to portray the sequence of the text. On the 3rd day, interventionists led children in review games (such as clue word bingo), and they guided students to make stories that followed the text structure genre of the week. These were brief comic-strip-like stories. On the 4th day, interventionists and students reviewed the clue words, then, to promote transfer, interventionists introduced a story that had the relevant text structure but that had no clue words and encouraged students to think about what words would best signal the text structure.

We hypothesized that TEXTS would improve struggling readers' listening comprehension and academic knowledge. We conjectured that with children's improved understanding of expository text structure they would be better able to understand more complex language structures and would gain academic knowledge through reading expository text.

The current study

This study was part of a larger study on the efficacy of the CTT component interventions (Connor et al., 2014) from prekindergarten through fourth grade (Phillips et al., 2018) with more than 8,000 children screened and more than 3,600 children participating in the studies. This study details the third- and fourth-grade findings. The overarching aim of the current study was to test aspects of the component model of reading for understanding—that improving third and fourth graders' key component skills would also improve their reading comprehension. We also conjectured that explicit interventions for these component skills would be more effective for children with weaker language and literacy skills, many of whom attend schools serving a higher proportion of children living in poverty. Thus, the following research questions guided this study:

  (1) To what extent do the component interventions impact the linguistic and cognitive skills they were designed to improve?
  (2) To what extent are there Child Characteristic × Treatment interaction effects. Specifically, are the interventions generally more effective for children with weaker language and decoding skills?
  (3) To what extent do the effects of the interventions also improve reading comprehension and other key component skills?)

Method

Participants

Overall, the final sample included 338 third- and 307 fourth-grade students attending 33 and 31 schools, respectively. Students in the study came from 135 third-grade classrooms and 115 fourth-grade classrooms. Participants were recruited from schools where at least 40% of the children were eligible for the NSLP, 52% of the students were girls, 39% were African American, 53% were White, 3% were multiracial, and the remaining children belonged to other ethnicities. On average, the third graders were 8.8 years of age ($SD = .50$) and the fourth graders were 9.8 years of age ($SD = .50$) at the time of the initial screening at the beginning of the school year. A CONSORT map is provided in Figure 1, which details the procedure for selecting students and inclusion criteria (i.e., vocabulary Standard Score < 98). Parent consent was obtained and screening completed for 1,106 third graders and 1,052 fourth graders. Overall attrition was 13%.

Measures

Detailed descriptions of the measures are provided in Table 1. We included multiple measures across three domains: vocabulary, syntax/listening comprehension (Lonigan & Milburn, 2017), and comprehension monitoring. We also included two measures of reading comprehension and measures of word reading. The

CONSORT Flow Diagram

COMPREHENSION TOOLS FOR TEACHERS (CTT) COMPONENT INTERVENTIONS STUDY
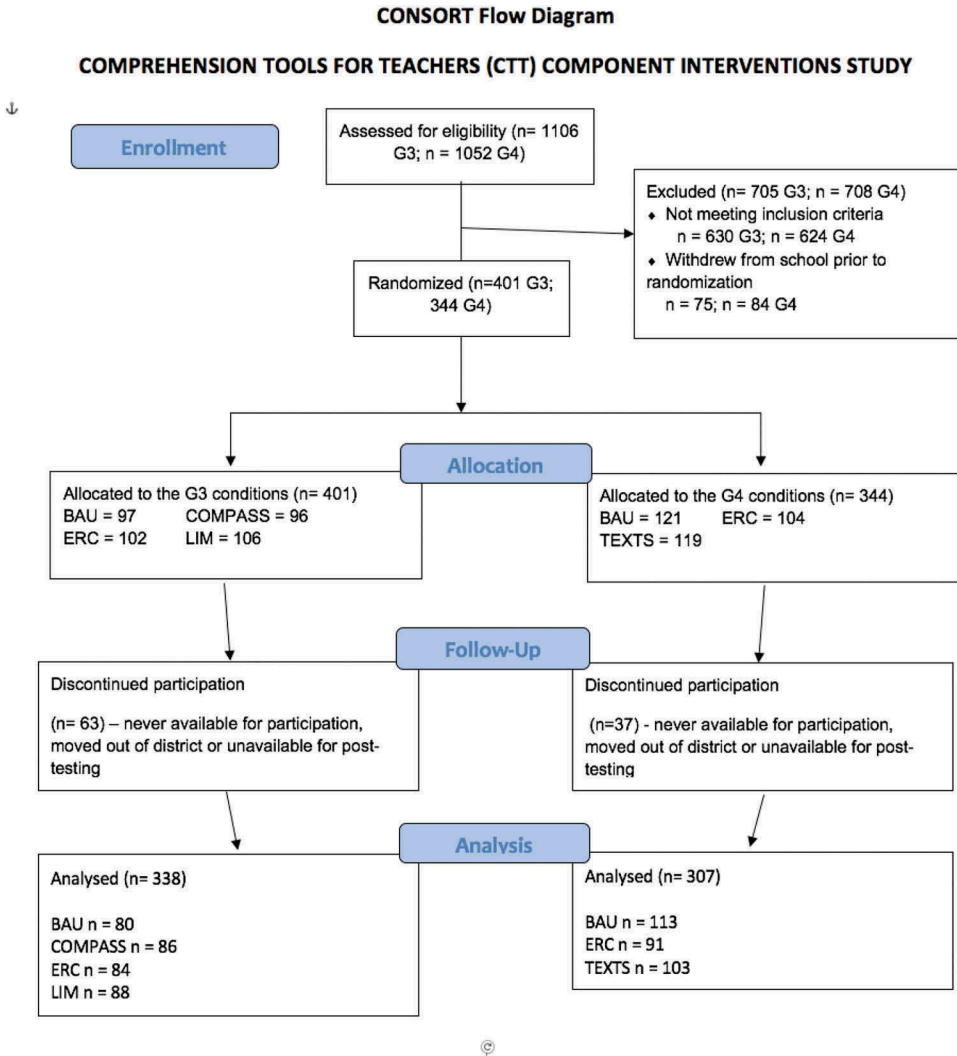
Figure 1. CONSORT Flow diagram of the Comparative Efficacy Study for third and fourth graders. Of them, 401 third graders and 344 fourth graders met the criteria for entrance into the study, which was a vocabulary standard score ($M = 100$, $SD = 15$) falling below 98 on the Expressive One Word Picture Vocabulary Test (Gardner, 1990). These students were then randomized to conditions. Of these, 63 third graders and 37 fourth graders left the study because they moved out of the district after randomization or were otherwise unavailable for posttesting (attrition of 16% and 11%, respectively).

latter were included to test for Child × Treatment interaction effects; no intervention specifically targeted fluent word reading. Measures were administered to all students in a quiet place in their school. Group-administered assessments were conducted in the students' classroom.

Procedures

Assignment to intervention condition

Because it was not possible to have every intervention group pair represented in each grade in every school, an incomplete-random-blocks design was used to assign students within schools to intervention conditions. Blocks representing each pairing of intervention conditions were randomized to school and order of use within each grade with the restriction that each block was roughly equally represented across schools.

Table 1. Measures Used in the Comparative Efficacy Study for Third and Fourth Graders.

Name and Citation	Description	Reliability
Vocabulary		
Expressive One Word Picture Vocabulary Test, Fourth Edition (EOWPVT) (Brownell, 2010)	The EOWPVT–4 assesses students' vocabulary through their ability to verbally label, with one word, objects, actions, and concepts when presented with color illustrations. Pictures and words become increasingly unfamiliar.	Normed for ages 2–18 years, internal consistency for ages 3–10 is acceptable (e.g., Cronbach's α range = .93–.97 and split-half reliability coefficients, corrected for the full length of the test range .96–.98).
CELF–4 Expressive Vocabulary (Semel, Wiig, & Secord, 2003)	This assessment evaluates the student's vocabulary through their ability to name illustrations of people, objects, and actions, including referential naming.	Normed for ages 5–9 years, reliability is acceptable with internal consistency reliability coefficient alphas for ages 5–9:11 years ranging from α = .80–.85; Test–retest stability coefficients (Corrected) for ages 6– 9:11 years ranging from .87 to .91.
Woodcock–Johnson–III Academic Knowledge (Woodcock, McGrew, & Mather, 2001)	Assesses students' academic knowledge in science, social studies, and the humanities by asking questions with picture support for some items. Students provide verbal responses to the questions.	The assessment has a median reliability of .88 in the 5–19 year range.
Syntax/Listening Comprehension		
Comprehensive Assessment of Spoken Language–Syntax Construction (Carrow-Woolfolk, 2008).	This assessment captures students' morphosyntactic abilities, asking them to formulate and express sentences using a variety of morphosyntactic rules.	Normed for ages 3–12 years, internal consistency using Rasch split-half method (odd/even) for ages 3–10 years is .73–.87; test–retest reliability for Syntax Construction (corrected) is 0.74 for third and fourth graders.
CELF Concepts and Following Directions (Semel et al., 2003)	Assesses listening comprehension by asking students to interpret and follow directions of increasing length and complexity (names, characteristics, and order of mention) using logical operations (e.g., before/after, tallest,, etc.).	Cronbach's alpha for ages 5–10:11 years ranges .73–.92.
Oral and Written Language Scales–Listening Comprehension scale (Carrow-Woolfolk, 1995)	Assesses students' ability to listen to and comprehend spoken language. Items are presented verbally and students point to the one of four pictures that best represents the item.	Split-half reliability is (α) .96–.98.
Test of Narrative Language Skills (Gillam & Pearson, 2004)	Measures students' ability to answer literal and inferential comprehension questions about narrative text	Cronbach's α = .74
Comprehension Monitoring		
Inconsistency Detection (Kim, 2017; Kim & Phillips, 2014)	This researcher-developed task is designed to assess students' comprehension monitoring during listening comprehension. Students listen to short scenarios with one inconsistent sentence, which the student is challenged to identify	Cronbach's α = .89;
Reading Comprehension Measures		
Test of Silent Reading Efficiency and Comprehension (Wagner, Torgesen, Rashotte, & Pearson, 2010)	Assesses the efficiency (i.e., speed and accuracy) with which student can read connected text silently for comprehension; Students are instructed: I want you to read some sentences and decide whether the answer is "yes" or "no." Look at the sample sentence. It says, "A cow is an animal." Is that true? (Pause for response) Because the answer is "yes," you should circle "yes" in the box.	Alternate-form (immediate administration) reliability coefficients is acceptable and range from .82 to .96 for Grades 3–5.
Gates MacGinitie Reading Test (MacGinitie & MacGinitie, 2006)	The test is group administered and evaluates students' abilities to understand extended written text. Students read passages and answer multiple-choice questions of increasing complexity.	The published reliability for this test is .96.

(*Continued*)

Name and Citation	Description	Reliability
Word Reading		
Woodcock–Johnson–III Letter Word Identification (Woodcock et al., 2001)	Measures students' ability to read printed letters and printed words correctly by reading lists of increasingly complex and unfamiliar words.	The assessment has a median reliability of .91 in the age range of 5–19 years.
The Test of Word Reading Efficiency–Second Edition (Torgesen, Wagner, & Rashotte, 1999)	This is a measure of an individual's ability to pronounce printed words (Sight Word Efficiency) and phonemically regular nonwords (Phonemic Decoding Efficiency) accurately and fluently. These timed tests are normed for persons 6–24 years of age.	Given that Cronbach's alpha and Split-half coefficients are considered inappropriate for timed tests, alternate forms reliability procedures for students 6– 10 years of age ranged .93–.97. Further, test–retest reliability for Forms A and B for 6–18 years of age ranged r = .84–.97.

Note. CELF = Clinical Evaluation of Language Fundamentals.

Children who qualified for the study within each grade were ordered by their score on the qualifying measure in groups of five to nine children (i.e., the minimum and maximum number of children that could be formed into two small-group instructional units) and randomized within block to one of the intervention conditions represented by their assignment block. Information about professional development and attendance is provided in the supplementary online materials.

Fidelity results

In general, CTT interventions in these two grades were delivered with strong fidelity (i.e., adherence and quality) with the highest possible rating equal to 5. In third grade, the average ratings for COMPASS were 4.21 (SD = .77) for adherence and 4.41 (SD = .64) for quality. For Language in Motion, average ratings were 4.20 (SD = .62) for adherence and 3.84 (SD = .77) for quality. Likewise, Enacted RC in third grade demonstrated averages of 4.61 (SD = .58) and 4.62 (SD = .53) for adherence and quality, respectively. In fourth grade, Enacted RC's average ratings were 4.71 (SD =.49) for adherence and 4.71 (SD = .50) for quality. Average ratings for TEXTS were 4.70 (SD = .36) and 4.93 (SD = .20) for adherence and quality, respectively.

Intervention procedures

Each component intervention was provided 4 days/week for 30 min. Lessons were implemented in small groups of between four and five students by a trained interventionist who was part of the research team. Each intervention lasted between 10 and 12 weeks.

Business-as-usual instruction

Based on brief observations of classrooms from which children were selected, instruction was of generally good quality and followed the districts' core literacy curriculum. Cores observed included Treasures, Wonders, Open Court Imagine, and Journeys. These core literacy curricula were approved for use in districts by the state of Florida at the time of the study. Of note, instructional focus in third and fourth grades was on reading comprehension, strategies, discussions about texts, building vocabulary in context, writing, and decoding/encoding. Based on our informal observations and discussions with teachers, as well as a survey we administered in the spring, it is unlikely that teachers provided the intensive focus provided by the CTT component interventions.

Results

Preliminary analyses

Descriptive statistics for standardized measures at pretest by intervention group are shown in Table S.1 for third-grade children and in Table S.2 for fourth-grade children in the supplementary

online materials. Correlations are also provided in Tables S.1 and S.2. As expected, all of the measures were moderately correlated with one another and, notably, the vocabulary, syntax/listening comprehension, and comprehension monitoring (Inconsistency Detection) measures were correlated with both reading comprehension measures, although none of the correlations were large.

Differences at baseline

Mixed models were used to determine if there were any differences between intervention groups at pretest. For third grade, on both the Sight-Word Efficiency and the Phonological-Decoding Efficiency subtests of the Test of Word Reading Efficiency (TOWRE), the Language in Motion group scored higher than the business-as-usual (BAU) group ($p = .02$). For fourth grade, there was one statistically significant effect. On the Concepts and Following Directions subtest of the Clinical Evaluation of Language Fundamentals (CELF), the TEXTS group scored higher than the BAU group ($p < .03$).

Overview of impact analysis

Our focus was on the two-group contrasts between the three (third grade) or two (fourth grade) experimenter-designed and delivered small-group interventions and the BAU group of children who received their regular classroom-based instruction only. Children were nested within schools and assignment block; therefore, all analyses used mixed models, treating these two factors as random effects. For some outcomes, one or both of these random effects resulted in zero or negative residual variance, indicating that the nesting variable did not account for variance in the outcome. Random effects were removed from the models sequentially to eliminate zero or negative residual variance in the final model. All models included children's ages, their raw scores on the Expressive One Word Picture Vocabulary Test (EOWPVT) at pretest, and their scores at pretest on the specific outcome variable in an analysis. The homogeneity of regression assumption of analysis of covariance was evaluated by including terms representing the Treatment Group × Covariate interaction for all covariates, and nonsignificant interaction terms were removed from the final model. Follow-up analyses were conducted to determine the size of the two-group comparisons at ± 1 SD of the covariate for any significant Treatment Group × Covariate interaction. These were modeled results and not subgroup comparisons. In these follow-up analyses, all other covariates and any of the other Treatment Group × Covariate interactions that were statistically significant were retained in the models.

We conducted multiple comparisons using a common control group; consequently, the danger of Type I error (finding an effect by chance) was inflated. At the same time, we were examining whether relatively brief, albeit intensive, interventions might have a positive effect on fairly insensitive standardized measures; thus, we also had the danger of Type II error (failing to identify a significant effect). We mitigated both types of error to the extent possible. Following the standards of the What Works Clearinghouse (n.d.), we applied the Benjamini–Hochberg (BH) correction (Benjamini & Hochberg, 1995) to outcomes within a domain (i.e., vocabulary [EOWPVT, CELF–Expressive Vocabulary, Woodcock–Johnson–II Academic Knowledge], syntax/listening comprehension [CELF-S, Oral and Written Language Scales, CELF–Concepts and Following Directions, Test of Narrative Language Skills], reading comprehension [Gates, Test of Silent Reading Efficiency and Comprehension], decoding [Woodcock–Johnson Letter–Word Identification subtest, TOWRE–Sight Word Efficiency, TOWRE–Phonemic Decoding Efficiency]; Lonigan, Burgess, & Anthony, 2000; Lonigan & Milburn, 2017). The BH correction controls the false-discovery rate instead of familywise error; therefore, it is less conservative than the Boneferroni method, limiting Type II error while providing adequate protection against Type I error. We computed BH corrected significance levels at both the conventional level of statistical significance (i.e., $p < .05$) and to detect marginal effects (i.e., $p < .10$). In the tables, we report the actual p values for all contrasts with p values of $p < .10$; however, only those that achieved significance following BH correction are marked as significant (*) or

marginally significant (~). Finally, we considered whether the effect sizes were educationally meaningful following Hill, Bloome, Black, and Lipsey (2008), where educationally meaningful effect sizes for third and fourth graders range from 0.24 to 0.48 with a mean of 0.36.

Impacts on outcomes for third-grade children

Overall contrasts

Posttest scores, adjusted for covariates in the final model, and effect sizes (Hedges's *g*) for the two-group comparisons for third-grade children are shown in Table 2. Of note, once the BH correction was applied, none of the contrasts were significant. However, some of the effects can be considered educationally meaningful. For example, for COMPASS, there was a treatment effect (*g*) of .31 on the Inconsistency Detection Task (i.e., comprehension monitoring), as hypothesized. For Enacted RC, there was an educationally meaningful effect of treatment (.33) on the EOWPVT (i.e., vocabulary), with children who received the intervention scoring higher than children in the BAU condition.

Follow-up contrasts: language outcomes

There were Intervention Condition × Covariate interactions, indicating that the effect of the intervention was not consistent among students based on their incoming skills. Many of the potentially educationally meaningful effects were not significant after the BH correction (see Table 3 and Figure 2). Highlighting those effects that were educationally meaningful and significant, we found that, for example, the effect of Enacted RC on vocabulary (the EOWPVT) was stronger for children who had lower EOWPVT scores prior to the start of the intervention than for children who had higher EOWPVT scores prior to the start of the intervention; the effect size for children with low vocabulary scores was large (.68).

Table 2. Descriptive Statistics for Posttest Scores and Effect Sizes for Two-Group Comparisons for Third-Grade Children.

| | Intervention Groups | | | | | | | | Effect Sizes | | |
| | BAU | | COMPASS | | ERC | | LIM | | | | |
Outcome	Adj. *M*	(SD)	Adj. *M*	(SD)	Adj. *M*	(SD)	Adj. *M*	(SD)	Compass vs BAU (p)	ERC vs BAU (p)	LIM vs BAU (p)
EOWPVT	84.67	(10.51)	86.26	(12.62)	88.22	(11.18)	86.35	(11.13)	**.14**	**.33 (.012)**	**.15**
CELF EV	38.14	(6.49)	38.66	(6.12)	38.98	(5.78)	38.29	(5.93)	**.08**	**.14**	**.02**
CASL S	27.74	(5.82)	28.46	(6.14)	28.71	(5.45)	28.21	(4.80)	**.12**	.17	**.09**
CELF CFD	41.25	(6.77)	42.19	(6.32)	41.78	(6.47)	40.29	(6.74)	**.14**	.08	**−.14**
OWLS	84.72	(8.36)	85.09	(10.22)	85.53	(9.46)	85.49	(10.09)	**.04**	.09	**.08**
TNL	29.73	(3.81)	30.25	(4.31)	30.11	(4.02)	29.88	(3.74)	**.13**	.10	**.04**
WJ AcKnow	15.47	(1.51)	15.53	(1.53)	15.69	(1.57)	15.69	(1.24)	.04	**.14**	.16
Incon Detect	15.36	(2.88)	16.27	(2.90)	15.56	(3.08)	14.83	(3.11)	**.31 (.065)**	.07	−.18
TOSREC	22.10	(6.09)	22.37	(6.46)	22.35	(6.21)	22.33	(6.86)	.04	.04	.04
GATES	26.93	(8.44)	26.57	(9.22)	26.14	(8.46)	26.15	(7.81)	−.04	−.09	−.10
WJ LIWID	48.35	(5.78)	47.73	(6.38)	48.08	(6.12)	47.95	(7.79)	−.10	−.05	−.06
TOWRE SWE	61.08	(10.57)	61.43	(11.20)	60.55	(10.91)	62.83	(8.60)	.03	−.05	.18
TOWRE PDE	29.78	(11.01)	28.92	(11.50)	29.88	(11.96)	29.52	(10.56)	−.08	.01	−.02

Note. No significant effects after correcting for multiple comparisons. Effect sizes in bold are for constructs that were hypothesized to be impacted by the intervention. BAU = Business-as-Usual control; COMPASS = Comprehension Monitoring and Providing Awareness of Story Structure; ERC = Enacted Reading Comprehension; LIM = Language in Motion; EOWPVT = Expressive One Word Picture Vocabulary Test; CELF = Clinical Evaluation of Language Fundamentals; EV = Expressive Vocabulary; CASL S = Comprehensive Assessment of Spoken Language–Syntax Construction; CFD = Concepts and Following Directions; OWLS = Oral and Written Language Scales; TNL = Test of Narrative Language Skills; WJ AcKnow = Woodcock–Johnson Academic Knowledge; Incon Detect = Inconsistency Detection; TOSREC = Test of Silent Reading Efficiency and Comprehension; GATES = Gates MacGinitie Reading Test; WJ LWIWD = Woodcock–Johnson Letter–Word Identification subtest; TOWRE = Test of Word Reading Efficiency; SWE = Sight Word Efficiency; PDE = Phonemic Decoding Efficiency.

Table 3. Differential Impacts of Intervention Conditions by Level of Moderator (in Parentheses) for Language-Related Outcomes for Third-Grade Children

Outcome (Moderator)	Value of Moderator	Adj. M for Intervention Group				Effect Size		
		BAU	COMPASS	ERC	LIM	Compass vs. BAU	ERC vs. BAU	LIM vs. BAU
EOWPVT	−1 SD	77.50	77.63	84.85	78.20	.01	.68* (.0009)	.06
	M	84.67	86.27	88.22	86.36	.15	.33 (.012)	.16
	+1 SD	91.84	94.91	91.59	94.52	.28	−.02	.25
CELF EV (EOWPVT)	−1 SD	35.34	37.34	37.49	35.60	.33 (.047)	.35 (.049)	.04
	M	38.13	38.65	38.98	38.28	.08	.14	.02
	+1 SD	40.91	39.97	40.47	40.96	−.15	−.07	.01
CELF CFD (EOWPVT)	−1 SD	39.19	39.43	41.06	39.94	.04	.28	.11
	M	41.24	42.17	41.78	40.29	.14	.08	−.14
	+1 SD	43.29	44.92	42.49	40.64	.25	−.12	−.39* (.0002)
CELF CFD	−1 SD	36.53	41.56	39.14	36.73	.76* (.0002)	.39 (.054)	.03
	M	41.25	42.19	41.78	40.29	.14	.08	−.14
	+1 SD	45.97	42.81	44.42	43.85	−.48 (.015)	−.23	−.31 (.065)
TNL (Age)	−1 SD	29.42	31.60	30.75	30.12	.56* (.006)	.34 (.076)	.19
	M	29.73	30.23	30.10	29.88	.13	.09	.04
	+1 SD	30.04	28.86	29.46	29.65	−.30	−.15	−.10

Note. For ERC × EOWPVT effect at the mean, the Benjamini–Hochberg critical value was equal to the *p* value. Effect sizes in bold are for constructs that were hypothesized to be impacted by the intervention. BAU = Business-as-Usual control; ERC = Enacted Reading Comprehension; LIM = Language in Motion; EV = Expressive Vocabulary; CELF = Clinical Evaluation of Language Fundamentals; EOWPVT = Expressive One Word Picture Vocabulary Test; EV = Expressive Vocabulary; CFD = Concepts and Following Directions; TNL = Test of Narrative Language Skills.

For the Concepts and Following Directions subtest of the CELF (i.e., simple listening comprehension) there was a large effect (.76) of COMPASS for children who had lower scores on this measure prior to the start of the intervention, whereas there was no impact for children with average or higher scores on this measure prior to the start of the intervention. There was also a significant negative effect of Language in Motion on Concepts and Following Directions subtest of the CELF for children who had higher scores on the measure and on vocabulary prior to the start of the intervention. Finally, there was a significant positive and educationally meaningful effect of COMPASS on the Test of Narrative Language (i.e., listening comprehension task that requires deep comprehension) for the youngest children, but there was no effect for children who were average age or older.

Follow-up contrasts: Word reading outcomes
There was a significant positive effect of Language in Motion on the Sight Word Efficiency subtest of the TOWRE for children who scored lower on the measure prior to the intervention (see Table 4) but no impacts for children who scored average or higher on the measure prior to the intervention. In addition, there was a significant negative effect of Enacted RC on the Sight Word Efficiency subtest of the TOWRE for children who scored lower on the EOWPVT (vocabulary) prior to the intervention but no impacts for children who scored at the average level or higher on the EOWPVT prior to the intervention.

Impacts on outcomes for fourth-grade children

Overall contrasts
For fourth graders, for Enacted RC and TEXTS (see Table 5), the largest effect size was 0.25, which might be considered educationally meaningful for this age range. However, after the BH correction, it was not statistically significant.

Follow-up contrasts: Language outcomes
There were also Intervention × Covariate interactions for fourth graders (see Table 6 and Figure 3). Again, many of the potentially educationally meaningful effect sizes were not significant after the BH correction. We highlight here those that were. There were significant and educationally meaningful
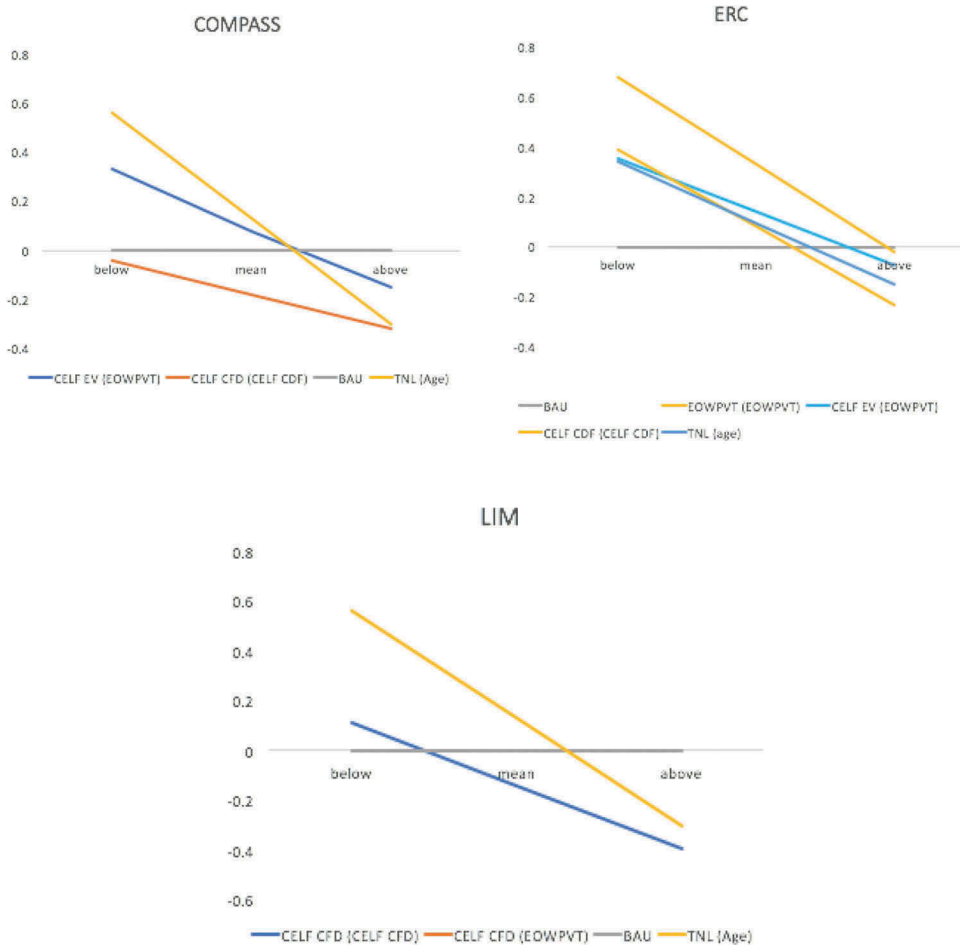
Figure 2. Third-Grade Child × Instruction Interaction posttest effect sizes (*g*) relative to the business-as-usual control (BAU) for Comprehension Monitoring and Providing Awareness of Story Structure (COMPASS), Enacted Reading Comprehension (ERC), and Language in Motion (LIM). *Note.* The moderator is the pretest, indicated in parentheses, falling 1 *SD* below the mean (below), at the mean (mean), or 1 *SD* above the mean (above). Clinical Evaluation of Language Fundamentals (CELF) Expressive Vocabulary (EV) and Expressive One Word Picture Vocabulary Test (EOWPVT) are vocabulary assessments, CELF Concepts and Following Directions (CFD) and Test of Narrative Language Skills (TNL) are measures of listening comprehension. Age is in years.

effects for both Enacted RC and TEXTS on Academic Knowledge for students who had lower scores on the measure prior to the start of the intervention. There was no significant effect for students with average scores on the measure prior to the start of the TEXTS and Enacted RC interventions. There was a significant negative effect of Enacted RC for students with strong academic knowledge skills at the beginning of the intervention.

Follow-up contrasts: Word reading outcomes
There was an educationally meaningful effect of Enacted RC on the Sight Word Efficiency subtest of the TOWRE for older children but no effects for children of average age or younger for their grade (see Table 7). In contrast, on the Letter-Word Identification, there was a positive effect of Enacted RC for younger children, and a negative effect for older children.

Table 4. Differential Impacts of Intervention Conditions by Level of Moderator (in Parentheses) for Code-Related Outcomes for Third-Grade Children.

Outcome (Moderator)	Value of Moderator	Adj. M for Intervention Group				Effect Size		
		BAU	COMPASS	ERC	LIM	Compass vs. BAU	ERC vs. BAU	LIM vs. BAU
	−1 SD	21.10	19.05	21.74	22.71	−.18	.06	.15
TOWRE PDE	M	29.78	28.92	29.88	29.52	−.07	.01	−.02
	+1 SD	38.47	38.80	38.02	36.34	.03	−.04	−.20
	−1 SD	61.68	61.84	58.17	63.90	.01	−.33	.23
TOWRE SWE (EOWPVT)	M	61.08	61.43	60.55	62.84	.03	−.05	.18
	+1 SD	60.48	61.01	62.92	61.77	.05	.23	.13
	−1 SD	52.58	54.24	51.94	57.05	.15	−.06	.47*
TOWRE SWE	M	61.08	61.43	60.56	62.84	.03	−.05	.18
	+1 SD	69.59	68.62	69.17	68.62	−.09	−.04	−.10

Note. BAU = Business-as-Usual control; COMPASS = Comprehension Monitoring and Providing Awareness of Story Structure; ERC = Enacted Reading Comprehension; LIM = Language in Motion; TOWRE = Test of Word Reading Efficiency; PDE = Phonemic Decoding Efficiency; SWE = Sight Word Efficiency; EOWPVT = Expressive One Word Picture Vocabulary Test.
*p significant after Benjamini-Hochberg correction.

Table 5. Descriptive Statistics for Posttest Scores and Effect Sizes for Two-Group Comparisons for Fourth-Grade Children.

	Intervention Groups						Effect Size	
	BAU		ERC		TEXTS			
	Adj. M	(SD)	Adj. M	(SD)	Adj. M	(SD)	ERC vs. BAU (p)	TEXTS vs. BAU (p)
EOWPVT	91.44	(10.82)	92.45	(11.44)	93.25	(11.44)	**.09**	.16
CELF EV	40.45	(5.68)	41.00	(5.95)	40.84	(5.70)	**.09**	.07
CASL S	31.76	(5.32)	31.00	(4.75)	32.29	(4.38)	−.15	**.11**
CELF CFD	44.03	(6.28)	44.17	(5.21)	44.52	(4.63)	.02	**.09**
OWLS	88.74	(10.13)	89.16	(8.65)	91.20	(9.50)	.04	**.25 (.049)**
Incon Detect	15.94	(2.54)	15.71	(2.77)	15.92	(2.70)	**−.09**	−.01
TNL	31.01	(3.63)	30.70	(3.29)	31.15	(3.38)	−.09	.04
WJ AcKnow	15.99	(1.87)	16.28	(1.42)	16.33	(1.58)	**.17**	**.20 (.067)**
TOSREC	28.08	(7.66)	28.38	(7.92)	27.56	(6.80)	.04	−.07
GATES	25.17	(8.54)	24.48	(8.86)	24.53	(7.86)	−.08	−.08
WJ LIWID	51.85	(6.73)	51.80	(5.72)	51.40	(5.93)	−.01	−.07
TOWRE SW	65.23	(12.00)	66.99	(9.63)	65.45	(9.85)	.16	.02
TOWRE NW	32.89	(11.19)	32.28	(11.34)	31.46	(10.60)	−.05	−.13

Note. No significant effects after applying the Benjamini-Hochberg correction. BAU = Business-as-Usual control; ERC = Enacted Reading Comprehension; TEXTS = Teaching Expository Text Structures; EOWPVT = Expressive One Word Picture Vocabulary Test; CELF = Clinical Evaluation of Language Fundamentals; EV = Expressive Vocabulary; CASL S = Comprehensive Assessment of Spoken Language–Syntax Construction; CFD = Concepts and Following Directions; OWLS = Oral and Written Language Scales; Incon Detect = Inconsistency Detection; TNL = Test of Narrative Language Skills; WJ AcKnow = Woodcock–Johnson Academic Knowledge; TOSREC = Test of Silent Reading Efficiency and Comprehension; GATES = Gates MacGinitie Reading Test; WJ LWIWD = Woodcock–Johnson Letter–Word Identification subtest; TOWRE = Test of Word Reading Efficiency; SW = Sight Word; NW = Non-word.

Discussion

This study was guided by three research questions. In the first, to what extent do the CTT component interventions impact the linguistic and metacognitive skills they were designed to improve, we found that for the most part, our relatively brief (10–12 weeks) but intensive (4 days/week) component interventions were generally not effective for many students in improving targeted skills assessed using standardized measures (see Figure 4 for a graphical representation of the findings). Although there were a few scattered educationally meaningful effect sizes (Hill et al., 2008), there were no main effects for any of the interventions once a BH correction was applied. Instead, the targeted interventions were effective for improving hypothesized skills but only for students with generally weaker incoming skills.

This finding emphasizes the importance of our second question: To what extent are there child Characteristic × Treatment interaction effects? Acknowledging that there may be other reasons for

Table 6. Differential Impacts of Intervention Conditions by Level of Moderator (in Parentheses) for Language-Related Outcomes for Fourth-Grade Children

Outcome (Moderator)	Value of Moderator	Adj. M for Intervention Group			Effect Size	
		BAU	ERC	TEXTS	ERC vs. BAU	TEXTS vs. BAU
CELF CFD	−1 SD	41.63	42.83	43.81	.21	**.40 (.030)**
	M	44.03	44.17	44.52	.02	**.09**
	+1 SD	46.43	45.52	45.24	−.16	**−.22**
WJ AcKnow	−1 SD	14.78	15.94	15.62	**.69* (.00002)**	**.48* (.003)**
	M	16.00	16.28	16.33	**.17**	**.19 (.067)**
	+1 SD	17.21	16.63	17.05	**−.35* (.029)**	**−.09**
TNL (Age)	−1 SD	32.25	30.83	31.80	−.41 (.029)	−.13
	M	31.02	30.70	31.16	−.09	.04
	+1 SD	29.80	30.58	30.52	.22	.20
TOSREC (EOWPVT)	−1 SD	26.43	27.71	28.12	.16	.23
	M	28.05	28.37	27.76	.04	−.04
	+1 SD	29.68	29.03	27.41	−.08	−.31 (.043)

Note. Effect sizes in bold are for constructs that were hypothesized to be impacted by the intervention. BAU = Business-as-Usual control; ERC = Enacted Reading Comprehension; TEXTS = Teaching Expository Text Structures; CELF CFD = Clinical Evaluation of Language Fundamentals Concepts and Following Directions; WJ AcKnow = Woodcock–Johnson Academic Knowledge; TNL = Test of Narrative Language Skills; TOSREC = Test of Silent Reading Efficiency and Comprehension; EOWPVT = Expressive One Word Picture Vocabulary Test.
*p significant after applying Benjamini-Hochberg correction. Actual *p* values, where *p* ≤ .10, are displayed.

Table 7. Differential Impacts of Intervention Conditions by Level of Moderator (in Parentheses) for Code-Related Outcomes for Fourth-Grade Children.

| Outcome (Moderator)	Value of Moderator	Adj. M for Intervention Group			Effect Size
		BAU	ERC	TEXTS	TEXTS vs. BAU
WJ LWID (Age)	−1 SD	52.60	51.80	52.33	−.13	−.04
	M	51.85	51.80	51.40	−.01	−.07
	+1 SD	51.09	51.79	50.46	.11	−.10
WJ LWID (EOWPVT)	−1 SD	51.23	52.50	51.47	.20*	.04
	M	51.85	51.80	51.40	−.01	−.07
	+1 SD	52.47	51.10	51.32	−.22*	−.18[†]
TOWRE SWE (EOWPVT)	−1 SD	64.23	67.07	66.84	.26[†]	.24[†]
	M	65.24	66.99	65.44	.16	.02
	+1 SD	66.26	66.91	64.04	.06	−.20
TOWRE SWE (Age)	−1 SD	67.41	66.36	66.14	−.10	−.11
	M	65.24	66.99	65.46	.16	.02
	+1 SD	63.07	67.61	64.78	.41***	.15

Note. BAU = Business-as-Usual control; ERC = Enacted Reading Comprehension; TEXTS = Teaching Expository Text Structures; WJ LWID = Woodcock–Johnson Letter–Word Identification subtest; EOWPVT = Expressive One Word Picture Vocabulary Test; TOWRE = Test of Word Reading Efficiency; SWE = Sight Word Efficiency.
[†]*p* < .10. *p < .05. ***p < .001.

the moderating effects besides the CTT interventions, the answer to this question proved to be complex but promising (see Figures 3 and 4). Keeping in mind that all of the students in the study scored below mean age/grade expectations on the vocabulary assessment (standard score < 98), the Child × Instruction interactions generally followed our hypotheses (see Figure 4). For example, as anticipated, COMPASS students, compared to the BAU students, had stronger listening comprehension, but this was the case only for students with weaker initial skills in third grade. Plus, COMPASS predicted stronger comprehension of literate language for younger children in third grade (i.e., 1 SD below mean) with no effect on older third graders. We wonder whether younger third graders are still developing comprehension of literate language, and thus key elements of the COMPASS lessons supported this development. In contrast, these skills may have been already developed for older children, so COMPASS had no effect on comprehension of literate language for older third graders. We saw a similar trend for Enacted RC, with positive effects for younger students and null to negative effects for older students through fourth grade. Enacted RC was also more effective in
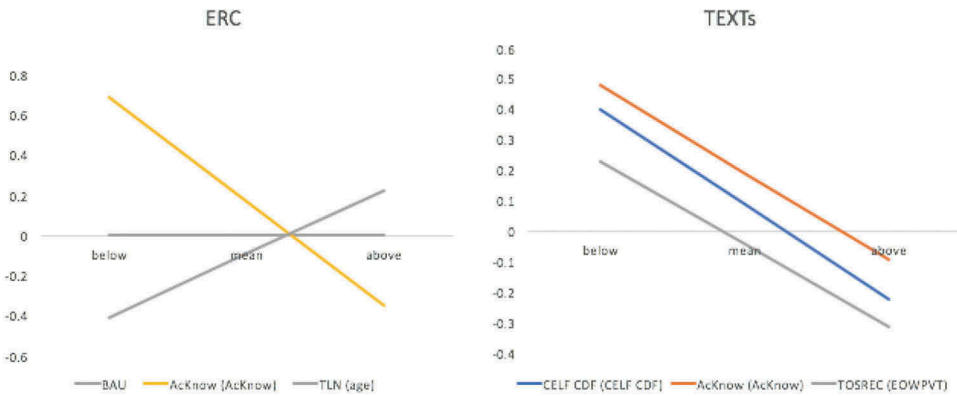
Figure 3. Fourth-Grade Child × Instruction Interaction posttest effect sizes (*g*) relative to the business-as-usual control (BAU) for Enacted Reading Comprehension (ERC) and Teaching Expository Text Structure (TEXTS). *Note.* The moderator is the pretest, indicated in parentheses, falling 1 *SD* below the mean (below), at the mean (mean), or 1 *SD* above the mean (above). Academic Knowledge (AcKnow) is a measure of academic knowledge; Clinical Evaluation of Language Fundamentals (CELF) is a measure of listening comprehension.
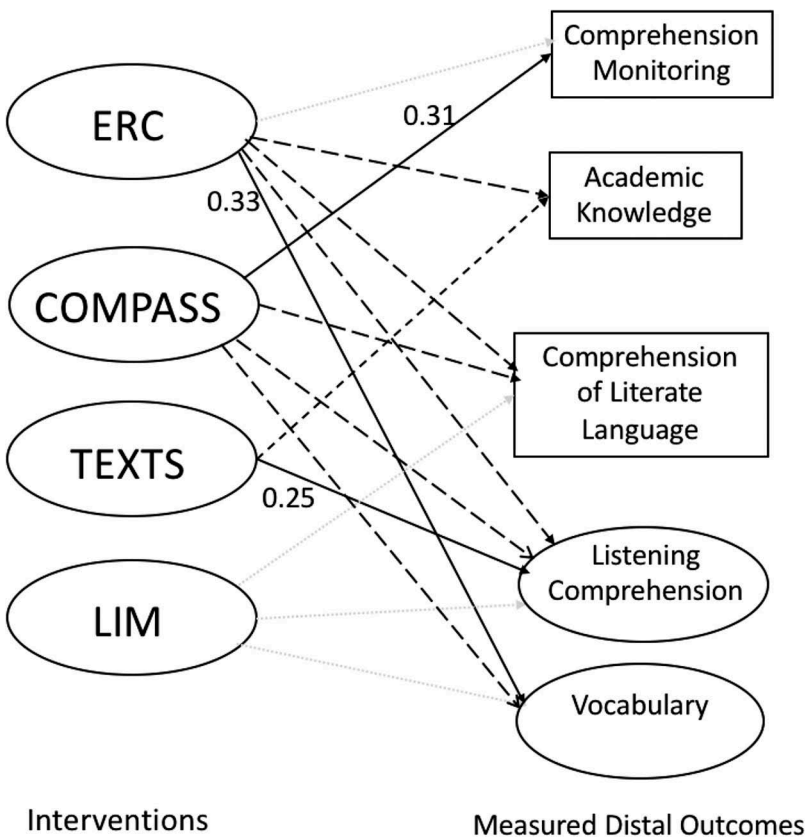
Figure 4. Graphical presentation of results for third and fourth grade combined. *Note.* Solid lines are educationally meaningful (*g* > 0.24) but not significant effects, long dash lines are interaction effects (i.e., the intervention worked for some children), and light dotted lines indicate the effect was not significantly different from 0. Coefficients represent main effect sizes (*g*) for either third or fourth grade. ERC = Enacted Reading Comprehension; LIM = Language in Motion.

improving vocabulary for students with weaker vocabulary skills compared to students with stronger vocabulary skills. Patterns of Child × Instruction interactions for fourth grade followed predicted patterns: TEXTS and Enacted RC were more effective in improving academic knowledge for students with weaker initial skills but not for students with typical and stronger skills. TEXTS was also effective in improving listening comprehension (CELF) for students with weaker initial skills but not for students with stronger skills.

With regard to word reading, there were unexpected Child × Instruction effects of the Language in Motion and Enacted RC interventions on silent word reading. Language in Motion (third grade only) had a positive treatment effect on silent word reading for children with weaker word reading skills. For older fourth graders participating in Enacted RC, there was a positive effect on silent word reading compared to the BAU group. We conjecture that this may have been related to the amount of reading required of the children. For example, in Enacted RC children read the entire chapter book *A Single Shard* (Park, 2001).

These consistent findings of Child × Instruction interaction effects in the moderation analyses, even for a somewhat homogeneous group (i.e., vocabulary < 98 standard score), are notable. If we consider only the children with more severely delayed language skills (i.e., > 1 *SD* below the sample mean), the pattern of findings remains similar: The component interventions were generally effective in improving targeted skills for students with weaker skills. At the same time, some of the interventions had negative effects for children with stronger skills, that is, we might assume that they would have been better off staying in the classroom rather than participating in the intervention. The causes for these unexpected results are unclear, but it appears that the targeted skills were not matched for the students' performance levels for the given outcomes. Our understanding of Child × Instruction interactions and how to personalize (or individualize or differentiate) the instruction provided is improving but still has a ways to go (Connor & Morrison, 2016).

Ultimately, the component interventions were designed to improve reading comprehension and other key component skills (our third research question). Unfortunately, the answer to this question was null; none of the interventions had an effect on either measure of reading comprehension, and any educationally meaningful treatment effects observed were generally confined to hypothesized targeted skills. Overall, the effects on key linguistic and metacognitive components of comprehension did not immediately generalize to stronger reading comprehension skills. Thus, taken with our findings for our first research question, we found that any learning was highly specific to what was taught, were effective only for some children, and did not generally transfer to other skills, including to reading comprehension. Limited generalization to other skills is consistent with evidence from other areas of research showing that training on specific skills often does not result in broader impacts on performance (e.g., "brain training" games improve specific game-related skills but do not provide general benefits to cognitive performance; Simons et al., 2016). Specificity of learning is an important concept with a long history in psychology (e.g., Woodworth & Thorndike, 1901) yet one we are likely to overlook as we design and test interventions and theories.

Results for the research questions together also suggest that there is limited support for component intervention models of reading comprehension, at least with our current intervention design (e.g., 10–12 weeks of intensive instruction on specific components). Whereas the evidence is overwhelming that the targeted skills are associated with reading comprehension (Oakhill & Cain, 2007; Hulme & Snowling, 2011; Nation, Clarke, Marshall, & Durand, 2004) and all of the measures in this study were positively correlated, it was not the case that improving the components skills in a targeted and specific way improved reading comprehension. New theoretical models are needed, and it was these findings and others that led us to the development of the lattice model of reading comprehension (Connor, 2016b; Connor et al., 2014), to which we return.

These findings also may help to explain some of the more disappointing findings regarding multi-tiered systems of support, which attempt to address individual student differences through a tiered system, with benchmarks set for when students receive more intensive Tier 2 and 3 interventions. Again, we consider the interventions presented here to be Tier 2 interventions, where Tier 1 is

general classroom instruction. A large national study, using regression discontinuity, revealed null and negative effects of placement in Tier 2 interventions for students who just made or just missed the schools' benchmark for placement in Tier 2 (Balu et al., 2015). In reviewing the participating schools' benchmarks, we noted that many schools set their benchmark at the 40th percentile on standardized assessments. This is comparable to our benchmark for inclusion in this study. Yet we found that the interventions were generally more effective for students with much weaker skills (closer to the 20th percentile). More careful consideration of multitiered systems of support benchmarks is warranted, as is thinking about how to help general education teachers personalize (differentiate, individual) both Tier 1 classroom instruction and Tier 2 interventions so that we can better serve children who bring diverse constellations of skills and backgrounds to the classroom and preserve intervention resources for children who need it most.

Limitations

Several limitations should be considered when interpreting these results. First, we conducted many analyses and applied a BH correction. Still, some effects might be by chance. The CTT component interventions were more likely to be effective for the component skills targeted, particularly for students with weaker incoming skills. For example, COMPASS, with its emphasis on comprehension monitoring and text structure via story retell, had educationally meaningful effects on inconsistency detection and interaction effects on listening comprehension, vocabulary, and comprehension of literate language relative to the BAU group. A second limitation is that, with one exception, no researcher-developed proximal measures were used in this study because the evaluation protocol was already lengthy. The decision was made when developing the study to use only well-validated standardized assessments. To be included in this study, each of the interventions had to demonstrate strong effects on proximal measures (Al Otaiba, Connor, & Crowe, in press; Connor et al., 2014; Kaschak et al., 2017; Kim & Phillips, 2014; Phillips et al., 2016). Finally, the reading comprehension measures, and some of the language measures we selected, may not have been sensitive to the language structures we targeted. As can be seen in Tables S.1 and S.2, none of the correlations between the outcome measures were moderate or large. We saw similar correlations with the Test of Silent Reading Efficiency and Comprehension.

Support for our original component model of reading was limited because, even when the intervention improved component skills, this did not generalize to improved reading comprehension. Improved skills also did not generalize to other language and metacognitive processes. However, other aspects of the model were supported; explicitly teaching language and metacognitive processes to students appeared to be somewhat effective in improving the targeted component skills for those students with weaker skills.

Further belying the component model, in other studies conducted by Reading for Understanding researchers, oral language appears to comprise two highly correlated factors—a semantic or vocabulary factor and a syntax/listening comprehension (or higher order) factor (Language and Reading Research Consortium, 2015; Lonigan & Milburn, 2017). Taking the findings of this study and the language-factor studies together, we conjecture that a more complete model of reading comprehension should consider reciprocating effects among text-specific, linguistic, social, and cognitive factors, which interact with instruction (i.e., Child × Instruction interactions) to impact children's reading for understanding (i.e., reading comprehension). Improving reading comprehension, in turn, influences children's developing text-specific, linguistic, social, and cognitive skills. There is emerging support for this model (Connor, 2016b; Connor et al., 2014, 2016). If one envisions this lattice of processes, instruction, and developing skills as a taut net, then improving only one component skill would be like plucking just one node of the net—the node would spring back to align with the rest of the net. Perhaps improving multiple nodes of the net simultaneously (e.g., multiple cognitive and linguistic interventions) might prevent the snapping-back action and, instead, improve the entire network of developing

skills and processes. That is, combining effective component interventions might be more effective than providing more focused single component interventions, particularly for children with weaker language and literacy skills.

In summary, whereas the overall results of the study revealed more null than positive impacts of the CTT interventions on language and metacognitive skills related to reading comprehension, these results have implications for theory, future research, and practice. As has been found before, improving reading comprehension continues to prove more difficult than anticipated. Still, these short but intensive CTT Tier 2 interventions were effective in improving some of their targeted skills, for some children, as assessed by fairly insensitive standardized measures. One might argue that improving language and metacognitive skills, even if there is no direct effect on reading comprehension, is a worthwhile endeavor—particularly because most third and fourth graders, including those with weaker oral language skills, typically do not receive specific language and metacognitive interventions. Notably, another study (Clarke et al., 2010) did find that improving components of oral language positively impacted reading comprehension. The age range for this study was similar to ours, but the criteria for entry to the Clark et al. study was different; students were selected into the study if they had fairly strong decoding skills but weak reading comprehension skills. Perhaps if we had selected students using reading comprehension as well as vocabulary in our criteria, our results might have aligned with Clark et al. Moreover, Clark and colleagues' effective interventions included multiple components of oral language including vocabulary, figurative language, and spoken narrative, which again suggests that multicomponent interventions may be more effective than more targeted intervention of a few components. We are now testing this hypothesis in studies by combining component interventions for prekindergarten and kindergarten students and for second graders. Finally, it is clear that new and more complex theories of the acquisition of reading comprehension, which include instruction and intervention, are needed. Although the component models are highly appealing, the results of this and other studies suggest that they are missing important active ingredients and wrongly assume skills will transfer to other related skills. We have proposed the lattice model, which has emerging support. However, more research, particularly intervention research to develop and test new models, is needed if we are to help more students proficiently read for understanding.

Acknowledgments

Funding

References

Al Otaiba, S., Connor, C. M., & Crowe, E. C. (2017). Promise and feasibility of teaching expository text structure: A primary grade pilot study. *Reading and Writing*. doi:10.1007/s11145-017-9769-6

Baker, L. (1984). Children's effective use of multiple standards for evaluating their comprehension. *Journal of Educational Psychology*, *76*, 588–597. doi:10.1037/0022-0663.76.4.588

Baker, L. (2008). Metacognition in comprehension instruction: What we've learned since NRP. In *Comprehension instruction: Research-based best practices* (*2nd*., pp. 65–79). New York, NY: Guilford Press.

Baker, L., & Stein, N. (1981). The development of prose comprehension skills. In C. S & B. Hayes (Eds.), *Children's prose comprehension: Research and practice* (pp. 7–43). Newark, DE: International Reading Association.

Balu, R., Zhu, P., Doolittle, F., Schiller, E. P., Jenkins, J. R., & Gersten, R. (2015). *Evaluation of response to intervention practices for elementary school reading*. Washington DC. Retrieved from http://ies.ed.gov/ncee/pubs/20164000/pdf/20164000.pdf

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.

Benson, S. E. (2009). Understanding literate language: Developmental and clinical issues. *Contemporary Issues in Communication Science and Disorders*, 36, 174–178.

Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test - Fourth Edition*. Minneapolis MN: Pearson.

Caillies, S., & Le Sourn-Bissaoui, S. (2008). Children's understanding of idioms and theory of mind development. *Developmental Science*, 11(5), 703–711. doi:10.1111/j.1467-7687.2008.00720.x

Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 47–72). Mahwah, NJ: Lawrence Erlbaum Associates.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, 96(1), 31–42. doi:10.1037/0022-0663.96.1.31

Cain, K., Towse, A. S., & Knight, R. S. (2009). The development of idiom comprehension: An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology*, 102, 280–298. doi:10.1016/j.jecp.2008.08.001

Carrow-Woolfolk, E. (1995). *Oral and Written Language Scales*. Circle Pines, MN: American Guidance Services.

Carrow-Woolfolk, E. (2008). *Comprehensive Assessment of Spoken Language*. Los Angeles, CA: Western Psychological Services.

Catts, H. W., Fey, M., Zhang, X., & Tomblin, B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, 3(4), 331–361. doi:10.1207/s1532799xssr0304_2

Chall, J. S. (1996). *Stages of reading development* (2nd ed.). Orlando, FL: Harcourt Brace.

Chen, Q.-S. (2009). Metacomprehension monitoring and regulation in reading comprehension. *Acta Psychologica Sinica*, 41(8), 676–683. doi:http://dx.doi.org/10.3724/SP.J.1041.2009.00676

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating Children's Reading-Comprehension Difficulties. *Psychological Science*, 21(8), 1106–1116. doi:10.1177/0956797610375449

Connor, C. M. (Ed.). (2016a). *The cognitive development of reading and reading comprehension*. London, UK: Routledge.

Connor, C. M. (2016b). A lattice model of the development of reading comprehension. *Child Development Perspectives*, 10(4), 269–274. doi:10.1111/cdep.12200

Connor, C. M., Day, S. L., Phillips, B., Sparapani, N., Ingebrand, S. W., McLean, L., & Kaschak, M. P. (2016). Reciprocal Effects of Self-Regulation, Semantic Knowledge, and Reading Comprehension in Early Elementary School. Child Development, 87(6), 1813–1824. doi:10.1111/cdev.12570

Connor, C. M., & Morrison, F. J. (2016). Individualizing student instruction in reading implications for policy and practice. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 54–61. doi:10.1177/2372732215624931

Connor, C. M., Phillips, B., Kaschak, M., Apel, K., Kim, Y.-S., Al Otaiba, S., & Lonigan, C. (2014). Comprehension tools for teachers: Reading for understanding from prekindergarten through fourth grade. *Educational Psychology Review*, 26(3), 379–401. doi:10.1007/s10648-014-9267-1

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology,* 99(2), 311–325. doi:10.1037/0022-0663.99.2.311

Efklides, A., & Misailidi, P. (Eds.). (2010). *Trends and prospects in metacognition research*. New York, NY: Springer.

Elliot-Faust, D. J., & Pressley, M. (1986). How to teach comparison processing to increase children's short- and long-term listening comprehension monitoring. *Journal of Educational Psychology*, 78, 27–33. doi:10.1037/0022-0663.78.1.27

Fitzgerald, J. (1989). Research on stories: Implications for teachers. In K. D. Muth (Ed.), *Children's comprehension of text: Research into practice* (pp. 2–36). Newark, DE:: International Reading Association.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–Developmental inquiry. *American Psychologist*, 34(10), 906–911. doi:10.1037/0003-066X.34.10.906

Gardner, M. F. (1990). *Expressive One Word Picture Vocabulary Test, EOWPVT* (Revised ed.). Novato, CA: Academic Therapy Publications.

Gersten, R., & Dimino, J. A. (2006). RTI (Response to Intervention): Rethinking Special Education for Students with Reading Difficulties (Yet Again). *Reading Research Quarterly*, 41(1), 99–108. doi:10.1598/RRQ.41.1.5

Gillam, R. B., & Pearson, N. A. (2004). *Test of narrative language*. Austin TX: Pro-Ed.

Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology*, 32, 389–399. doi:10.1016/j.cedpsych.2006.03.001

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japunitich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Research*, 96(3), 424–436.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, 9, 558–565. doi:10.3758/BF03196313

Graesser, A., Golding, J. M., & Long, D. L. (1991). Narrative representation and comprehension. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 171–205). White Plains, NY: Longman.

Hill, C., Bloome, H., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172–177. doi:10.1111/j.1750-8606.2008.00061.x

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2(2), 127–160. doi:10.1007/BF00401799

Hulme, C., Nash, H. M., Gooch, D., Lervåg, A., & Snowling, M. J. (2015). The Foundations of Literacy Development in Children at Familial Risk of Dyslexia. *Psychological Science*, 26(12), 1877–1886. doi:10.1177/0956797615603702

Hulme, C., & Snowling, M. J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science*, 20(3), 139–142. doi:10.1177/0963721411408673

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27(2), 236–248. doi:10.1037/0012-1649.27.2.236

James-Burdumy, S., Mansfield, W., Deke, J., Carey, N., Lugo-Gil, J., Hershey, A., & Faddis, B. (2009). *Effectiveness of Selected Supplemental Reading Comprehension Interventions: Impacts on a First Cohort of Fifth-Grade Students. NCEE 2009-4032*. Retrieved from http://search.proquest.com/docview/61882547?accountid=4840

Joshi, R. M., & Aaron, P. G. (2000). THE COMPONENT MODEL OF READING: SIMPLE VIEW OF READING MADE A LITTLE MORE COMPLEX. *Reading Psychology*, 21(2), 85–97. doi:10.1080/02702710050084428

Kaschak, M. P., Connor, C. M., & Dombek, J. L. (2017). Enacted Reading Comprehension: Using bodily movement to aid the comprehension of abstract text content. *PLoS ONE*, 12(1), e0169711. doi:10.1371/journal.pone.0169711

Kim, Y.-S. (2015). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. *Child Development*, 86, 128–144. doi:10.1111/cdev.12293

Kim, Y.-S. (2016). Direct and mediated effects of language and cognitive skills on comprehension or oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology*, 141, 101–120. doi:10.1016/j.jecp.2015.08.003

Kim, Y.-S. (2017). Why the simple view of reading is not simplistic: Unpacking the simple view of reading using a direct and indirect effect model of reading (DIER). *Scientific Studies of Reading*, 21, 310–333. doi:10.1080/10888438.2017.1291643

Kim, Y.-S., & Phillips, B. (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly*, 49, 269–281. doi:10.1080/10888438.2015.1007375

Kim, Y.-S., & Phillips, B. (2016). 5 minutes a day: An exploratory study of improving comprehension monitoring for prekindergartners from low income families. *Topics in Language Disorder*, 36, 356–367. doi:10.1097/TLD.0000000000000103

Kim, Y.-S. G., & Pilcher, H. (2016). What is listening comprehension and what does it take to improve listening comprehension? In R. Schiff & M. Joshi (Eds.), Handbook of interventions in learning disabilities (pp. 159–174). New York: Springer.

Kim, Y.-S. G., & Wagner, R. K. (2015). Text (Oral) reading fluency as a construct in reading development: An investigation of its mediating role for children from Grades 1 to 4. *Scientific Studies of Reading*, 19, 224–242. doi:10.1080/10888438.2015.1007375

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge Unviersity Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Language and Reading Research Consortium. (2015). The dimensionality of language in young children. *Child Development*, 86(6), 1948–1965. doi:10.1111/cdev.12450

Locke, J. L. (1997). A Theory of Neurolinguistic Development. *Brain and Language*, 58, 265–326. doi:10.1006/brln.1997.1791

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. *Developmental Psychology*, 36, 596–613. doi:10.1037/0012-1649.36.5.596

Lonigan, C. J., & Milburn, T. F. (2017). Identifying the dimensionality of oral language skills of children with typical development in preschool through fifth grade. *Journal of Speech, Language, and Hearing Research*, 60(8), 2185–2198. doi:10.1044/2017_JSLHR-L-15-0402

MacGinitie, W. H., & MacGinitie, R. K. (2006). *Gates-MacGinitie Reading Tests* (4th ed.). Iowa City, IA: Houghton Mifflin.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, 80(3), 300–335. doi:10.3102/0034654310377087

Megherbi, H., & Ehrlich, M. F. (2005). Language impairment in less skilled comprehenders: The on-line processing of anaphoric pronouns in a listening situation. *Reading and Writing*, 18(7), 715–753. doi:10.1007/s11145-005-8131-6

Mellard, D. F., & Fall, E. (2012). Component Model of Reading Comprehension for Adult Education Participants. *Learning Disability Quarterly*, 35(1), 10–23. doi:10.1177/0731948711429197

Meyer, B. J. F., & Poon, L. W. (2001). Effects of structure strategy training and signaling on recall of text. *Journal of Educational Psychology*, 93, 141–159. doi:10.1037/0022-0663.93.1.141

Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development, 61*(3), 722–730.

Morrison, F. J., Bachman, H. J., & Connor, C. M. (2005). *Improving literacy in America: Guidelines from research*. New Haven, CT: Yale University Press.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairments?. *Journal of Speech, Language, and Hearing Research*, 47, 199–211. doi:10.1044/1092-4388(2004/017)

National Assessment Governing Board. (2007). *Reading framework fo the 2007 National Assessment of Educational Progress*. Washington DC: U.S. Department of Education.

Oakhill, J., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, 18, 443–468. doi:10.1080/01690960344000008

Park, L. S. (2001). *A single shard*. New York, NY: Houghton Mifflin Harcourt Publishing Co.

Phillips, B., Kim, Y.-S. G., Lonigan, C. J., Connor, C. M., Clancy, J., & Al Otaiba, S. (2018). Supporting language and literacy development with intensive small-group interventions: An early childhood efficacy study. Manuscript submitted for publication.

Pearson, B. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, 8, 317–344. doi:10.1016/0361-476X(83)90019-X

Phillips, B. M., Tabulda, G., Ingrole, S. A., Burris, P. W., Sedgwick, T. K., & Chen, S. (2016). Literate Language Intervention With High-Need Prekindergarten Children: A Randomized Trial. *Journal of Speech, Language, and Hearing Research*, 59(6), 1409–1420. doi:10.1044/2016_JSLHR-L-15-0155

Rand Study Group, & Snow, C. E. (2001). *Reading for understanding*. Santa Monica, CA: RAND Education and the Science and Technology Policy Institute.

Rapp, D. N., Van Den Broek, P., McMaster, K., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, 11(4), 312–389. doi:10.1080/10888430701530417

Ruffman, T. (1999). Children's understanding of logical inconsistency. *Child Development*, 70, 872–886. doi:10.1111/cdev.1999.70.issue-4

Scarborough, H. S. (1998). Early identification of children at risk for reading disablties: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–119). Timonium, MD: York Press, Inc.

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York, NY: Guilford Press.

Schwanenflugel, P. J., Fabricius, W. V., & Noyes, C. R. (1996). Developing organization of mental verbs: Evidence for the development of a constructivist theory of mind in middle childhood. *Cognitive Development, 11*, 265–294.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals (CELF)* (4th ed.). San Antonio, TX: PsychCorp.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do "brain-training" programs work?. *Psychological Science in the Public Interest*, 17, 103–186. doi:10.1177/1529100616661983

Skarakis-Doyle, E., Dempsey, L., & Lee, C. (2008). Identification of language comprehension impairments in preschool children. *Language, Speech, and Hearing Services in Schools*, 39, 54–65. doi:10.1044/0161-1461(2008/006)

Stevens, R. J., Van Meter, P., & Warcholak, N. D. (2010). The effects of explicitly teaching story structure to primary grade children. *Journal of Literacy Research*, 42, 159–198. doi:10.1080/10862961003796173

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934–947.

Tompkins, G. E., & McGee, L. M. (1989). Teaching repetition as a story structure. In K. D & Muth (Eds.), *Children's comprehension of text: Research into practice* (pp. 59–78). Newark, DE:: International Reading Association.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency: TOWRE*. Austin, TX: Pro-Ed.

Van Den Broek, P., Risden, K., Fletcher, C. R., & Thorlow, R. (1996). A "landscape view" of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Brtitton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165–187). Mahweh, NJ: Lawrence Erlbaum Associates.

Van der Lely, H. K. J., & Marshall, C. R. (2010). Assessing component language deficits in the early detection of reading difficulty risk. *Journal of Learning Disabilities, 43*, 357–368.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *TOSREC: Test of silent reading efficiency and comprehension.* Austin, TX: Pro-Ed.

What Works Clearinghouse. (n.d.) WWC procedures and standards handbook, version 4. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Williams, J. P., Hall, K. M., Lauer, K. D., Stafford, B., DeSisto, L. A., & deCani, J. (2005). Expository text comprehension in the primary grade classroom. *Journal of Educational Psychology*, *97*(4), 538–550. doi:10.1037/0022-0663.97.4.538

Williams, J. P., Stafford, K. B., Lauer, K. D., Hall, K. M., & Pollini, S. (2009). Embedding reading comprehension training in content-area instruction. *Journal of Educational Psychology*, *101*(1), 1–20. doi:10.1037/a0013152

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson–III Tests of Achievement*. Itasca, IL: Riverside.

Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, *8*, 247–261. doi:10.1037/h0074898

Zipoli, J., . R. (2017). Unraveling difficult sentences: Strategies to support reading comprehension. *Intervention in School and Clinic*, *52*, 218–227. doi:10.1177/1053451216659465

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135(1-11)*. doi:10.1037/0096-3445.135.1.1