

A Meta-Analysis on the Effects of Text Structure Instruction on Reading Comprehension in the Upper Elementary Grades

Suzanne T.M.

Bogaerds-Hazenberg

Jacqueline Evers-Vermeul

Huib van den Bergh

*Utrecht University, Utrecht,
The Netherlands*

ABSTRACT

In this meta-analysis, the authors synthesize results from 44 (quasi-) experimental studies on informational and narrative text structure interventions involving students in grades 4-6 in regular school settings. Findings show that text structure instruction had positive immediate effects on students' reading comprehension but that effect sizes varied largely across outcome measures: questions (Hedges' $g = 0.25$), summarization ($g = 0.57$), recall ($g = 0.37$), and knowledge about text structure ($g = 0.38$). However, students who received text structure instruction no longer outperformed control groups at delayed posttests. Content-related features, such as a focus on paragraph-level structure, active construction of graphic organizers, and teaching rule-based summarization techniques, moderated the effectiveness of text structure instruction, but these effects also varied across outcome measures. Instructional features moderated delayed effects: Interventions with opportunities for individual student practice resulted in higher delayed effects for comprehension questions. The authors argue that text structure instruction deserves a place in the primary school curriculum so the positive effects on reading will be maintained.

Good reading comprehension skills are crucial for understanding text and play a pivotal role in academic, social, and economic success (Oakhill, Cain, & Elbro, 2015; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). However, reading comprehension is a complex skill, requiring both fluent decoding abilities and good language proficiency (Gough & Tunmer, 1986; Scarborough, 2001), both of which need to be promoted through instruction (Bus & van IJzendoorn, 1999; Oakhill et al., 2015). Our aim in the current meta-analysis was to examine whether text structure instruction can successfully improve the reading comprehension of students in the upper elementary grades and to determine what content and instructional components are related to the best outcomes.

According to national standards, by the end of primary education, students should be able to understand simple narrative and informational texts, distinguish various genres, and learn from texts (e.g., Expertgroep Doorlopende Leerlijnen Taal en Rekenen, 2009; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Despite intensive instruction, a substantial number of elementary school students struggle with reading comprehension (Kendeou & van den

Broek, 2007). Comprehension problems especially arise when students enter fourth grade and have to make the transition from learning to read to reading to learn, an effect known as the fourth-grade slump (Chall & Jacobs, 1983).

One of the factors that contribute to poor text comprehension is readers' inability to perceive the meaningful relations between information units (e.g., events, facts, settings) in a text (van den Broek & Kremer, 2000). As a result, readers construct a representation of the textbase—propositions that are directly derived from the text at the sentence level—but fail to understand how these propositions are organized on a global level. According to the construction–integration model (Kintsch, 1988, 2013; van Dijk & Kintsch, 1983), it is precisely this (re)organization of propositional information at a global level and the successive integration with prior knowledge that are crucial for text comprehension (Stine-Morrow, Gagne, Morrow, & DeWall, 2004; van der Schoot, Horsley, & van Lieshout, 2010). The better the information is organized in mental schemata and elaborated with relevant prior knowledge, the more coherent the readers' situation model of the text is (Kintsch, 1988, 2013; Kintsch & van Dijk, 1978) and, hence, the better their understanding is.

When readers are sensitive to the hierarchical organization of information in texts, this can facilitate the construction of the situation model (Kintsch, 2013). Several rhetorical patterns in the organization of information appear in many texts, such as cause-and-effect and compare-and-contrast (Meyer, 1975). Pyle et al. (2017) defined *text structure* as “the organization of ideas, the relationship among the ideas, and the vocabulary used to convey meaning to the reader (Armbruster, 2004; Shanahan et al., 2010)” (p. 469). In short, text structure influences how readers read and how writers write (Jiang & Grabe, 2007) and, hence, not only describes the text itself but also characterizes readers' cognitive coherence representations (e.g., Meyer & Freedle, 1984; Sanders & Noordman, 2000).

Empirical research has shown that struggling readers typically fail to rely on text structure to guide their reading (Meyer, Brandt, & Bluth, 1980; Rapp et al., 2007). By contrast, proficient readers make active use of text structure to organize their memory for textual content; they attend to both the external physical organization of the text (e.g., headings, table of contents) and the internal structure of ideas for a better understanding (Anderson & Pearson, 1984; Kendeou & van den Broek, 2007; Meyer & Rice, 1984). They construct a higher order structure of text while reading, which “guides encoding, recall, and reproduction of the essential points of the text” (Armbruster, Anderson, & Ostertag, 1987, p. 332). Proficient readers' knowledge about genre or structure influences their expectations about the text and helps them better predict and organize textual content while reading (Zwaan, 1994). They are better able to ask relevant questions about the essential points of the text while reading, which helps them

monitor their comprehension more effectively (Gersten, Fuchs, Williams, & Baker, 2001). From an instructional perspective, explicit teaching of text structure therefore seems a fruitful way to help students better anticipate, predict, and monitor their understanding of a text (Ogle & Blachowicz, 2002).

These insights have inspired a wide variety of interventions aimed at teaching the internal structure of texts as a strategy to improve students' comprehension and recall of texts. Two meta-analyses have shown that text structure instruction is a promising approach for improving text comprehension for learners of various ages (i.e., kindergarten to high school) and abilities (i.e., with and without learning disabilities), with overall effect sizes of Hedges' $g = 0.56$ (Hebert, Bohaty, Nelson, & Brown, 2016) and Cohen's $d = 0.95$ (Pyle et al., 2017). In our meta-analysis, we refined these meta-analyses by narrowing the focus to students in the upper elementary grades, who are faced with increasing literacy demands as they transition from learning to read to reading to learn (Chall & Jacobs, 1983).

At the same time, we broadened the scope by including studies on both narrative and informational text structures, and interventions focused on narrow text structure instruction (i.e., recognizing structures). In addition, we included closely related studies addressing structure-based summarization training, paragraph-level structure (e.g., topic sentences), and graphic organizer (GO) instruction. Although these studies did not explicitly train students in naming specific text structures, they focused students' attention on top-level structures. Moreover, summarization techniques and GOs are often part of text structure interventions, as such activities can promote students' sensitivity to hierarchical discourse patterns in texts, which could facilitate situation model construction (Kintsch, 2013).

Although Pyle et al. (2017) pointed out that it is important to find out whether and how instructional features (e.g., including collaborative activities) moderate the effectiveness of these interventions, they did not include an analysis of such features. Therefore, in our meta-analysis, we also examined the moderating impact of instructional features such as teacher modeling and collaborative versus individual student practice. This fits in well with research suggesting that more attention should be paid to the instructional context of text structure instruction (Beerwinkle, Wijekumar, Walpole, & Aguis, 2018; Pyle et al., 2017; Turcotte, Giguère, & Godbout, 2015; Wijekumar, Beerwinkle, Harris, & Graham, 2019; Williams, 2018).

A growing body of evidence suggests that reading comprehension is not a unitary construct and that different comprehension tests (e.g., questions, recall, summarization, text structure knowledge) measure different aspects of the reading process (Keenan, Betjemann, & Olson, 2008; Nation & Snowling, 1997). Although previous meta-analyses generated smaller effects on standardized measures than on researcher-developed measures (Hebert et al., 2016; Pyle

et al., 2017) and showed that the largest effects were found on GO tasks (Pyle et al., 2017), these meta-analyses merged outcome measures when it came to moderator analyses of specific content-related variables. Therefore, it remains unclear whether, for instance, a focus on paragraph-level structure affects recall or only summarization.

In terms of maintenance, previous meta-analyses displayed positive delayed effects on reading achievement, although these effects were smaller and typically lacked consistency, and the median delay between immediate and delayed posttest was only seven days (Hebert et al., 2016). However, it is not yet clear how delayed effects are affected by intervention characteristics and if they vary per outcome measure. This is why we examined the impact of content-related and instructional features per outcome measure on immediate and delayed effects, which provides a valuable theoretical and methodological addition. In sum, in our meta-analysis, we evaluated the moderating effects of various content-related and instructional variables in interaction with various outcome measures on students' reading comprehension to refine and expand our knowledge of the ingredients that make text structure instruction worthwhile. Two research questions guided our meta-analysis:

1. What are the immediate and delayed effects of text structure instruction on students' text comprehension in grades 4–6, as measured by comprehension questions, recall, summarization, and knowledge of text structures?
2. How are these immediate and delayed effect sizes moderated by content-related and instructional features?

In the following sections, we provide an overview of the content-related and instructional features that were taken into account in this meta-analysis, thereby highlighting the state of the art of research on text structure interventions. These features are necessary to describe the categories of analysis and the specific issues that we investigated in our meta-analysis.

Content-Related Features of Text Structure Instruction

Reading instruction in the primary grades often starts with narrative texts, followed by an increasing number of expository texts in the upper elementary grades (Chall & Jacobs, 2003). In the 1970s, story grammar instruction was developed to aid students in their comprehension of stories (e.g., Hansen, 1978). However, it was soon discovered that students struggled most with expository text comprehension (Taylor & Beach, 1984), possibly because expository texts contain a high amount of specific vocabulary and many unfamiliar concepts and vary more in their

underlying text structure than stories typically do (Hiebert & Mesmer, 2013; Pyle et al., 2017). In addition, the reading curriculum in primary schools was strongly focused on narrative texts (Duke, 2000; Durkin, 1978; Moss & Newton, 2002), which resulted in limited exposure to expository texts. As a consequence, primary school students displayed fewer spontaneously developed intuitions about expository text structure (Goldman, 1997). Since the early 1980s, many interventions on narrative and expository text structures have been developed, which often consist of combinations of the following features:

- Structure recognition (e.g., identifying top-level and paragraph-level structures, story grammar)
- Structure visualization (e.g., story mapping, GOs)
- Structure-based summarization (e.g., hierarchical outlining, rule-based summarization)

In this section, we discuss these features, before turning to the instructional variables frequently found in text structure interventions.

Structure Recognition

Most text structure interventions focus on teaching students how to recognize the top-level structure of expository texts (e.g., compare-and-contrast, cause-and-effect). Research on this topic started in the mid-1970s and was strongly influenced by information-processing theories focusing on cognitive processes that affect storage and retrieval of information (Kelly, 2019). One pioneer study was carried out by Bartlett (1978), who found that a training in text structure recognition increased ninth graders' ability to identify a text's top-level structure and use it for recall. Text structure recognition seems to raise text comprehension, especially when multiple text structures are taught (Hebert et al., 2016; Pyle et al., 2017).

Meyer and Ray (2011) provided an excellent overview of interventions focused on structure recognition. Typically, these interventions consist of teaching questions that are answered (e.g., what are the differences between A and B?) and practice items in which students categorize short texts as belonging to one structure or another. Also, students learn about cue words or signaling words that frequently appear in these types of structures (e.g., *similar* or *likewise* in compare-and-contrast texts), as these words instruct readers in how to process an upcoming information segment and how to relate it to a previous one, thereby assisting them toward building coherent text representations (Sanders & Sporeen, 2007; van Silfhout, Evers-Vermeul, & Sanders, 2015). In most interventions, cue words are simply listed in booklets, highlighted or mentioned as characteristics of a specific text structure, but in some interventions, students actively highlight (Bohaty, 2015), annotate (Gentry, 2006; Short & Ryan, 1984), or write down cue words found in a text (Broer, Aarnoutse, Kieviet, & van Leeuwe, 2002).

Especially in the context of informational texts, structure recognition training can also be focused on the paragraph-level of the text, such as by teaching students how paragraphs are typically structured in topic sentences, supporting details, and concluding sentences. Often, students receive explicit instruction about the main idea or on topic sentences and then read a text and select for each paragraph the sentence or phrase that captures the most important information at the highest level (Broer et al., 2002; de Jou & Sperb, 2009; Vidal-Abarca, 1990), or students learn through teacher modeling how to invent a good summarizing phrase when there is no clear topic sentence (Braxton, 2009). Sometimes, students learn more than simply distinguishing between main ideas and details. For instance, Gentry (2006) taught students to make annotations in the margins of paragraphs, such as writing “Ex.” when the text discussed an example. In many of these interventions, paragraph-level instruction was combined with top-level structure recognition (i.e., instruction on the features of different structures).

In other studies, students learned about the blueprint of narrative texts, also called story grammar. This often involves teaching students how to identify the protagonists, their goals or problems, the actions, and the outcome (Gersten et al., 2001; Zwaan & Radvansky, 1998). Students typically receive instruction on the basic elements of a story and use this knowledge to analyze a short story (e.g., Idol & Croll, 1987). Research on these interventions in the upper elementary grades with typically developing students has been scarce, as many of these interventions are focused on younger students or students with learning difficulties (Gersten et al., 2001). In general, it seems that students with knowledge of story grammar are better able to make predictions about a text, recognize what information is crucial for the plot (Wolman, 1991), and recall more about the main story elements, such as the setting or protagonist (e.g., Hansen, 1978; Mandler & Johnson, 1977; Weaver & Dickinson, 1982).

In this meta-analysis, we examined the effects of text structure recognition on comprehension. In particular, we analyzed whether informational text structure instruction and narrative text structure instruction have similar effects and whether the number of text structures taught matters.

Structure Visualization

Another family of strategies for improving comprehension and recall is to teach students how to visualize the organization of main ideas via GOs. GOs can be used as visualizations of the hierarchical relations of textual information (Griffin, Malone, & Kameenui, 1995), in which relations between concepts are communicated through the visual placement of concepts relative to each other (Robinson & Molina, 2002). Some common GOs are Venn diagrams, matrices, knowledge maps, and tree diagrams (Manoli &

Papadopoulou, 2012), but not outlines and lists, as these miss a visual argument and are more textlike (Hoffman, 2010).

GO development probably started in the early 1970s in Tokyo, Japan, where Ishikawa (1971) developed so-called fishbone diagrams, early cause-and-effect GOs that were used to control the supply chain and manufacturing process in the shipbuilding industry. In educational contexts, GOs first became vogue as a variation on advance organizers (Barron, 1980). For a long time, GOs were used in prereading activities to activate and organize students' prior content knowledge (Moorf & Readence, 1984), but nowadays, GOs are also used as a visualization strategy during reading (Leutner, Leopold, & Sumfleth, 2009) or as postreading summarization activities. The improved access to digital resources has made computer-based GOs and concept-mapping software (e.g., Kidspiration, Webspiration) increasingly popular in schools (Ciullo & Reutebuch, 2013; Smith & Okolo, 2010). Even so, most studies on structure visualization have been centered on print-based modalities.

GOs constitute a valuable way to enhance comprehension because they can show the main information of a text at a glance and simultaneously clarify relations between these ideas (Jones, Pierce, & Hunter, 1988). Structure visualization in GOs constitutes a major way to make students aware of text structure, as they provide a visual map of the structure (Jiang & Grabe, 2007; Manoli & Papadopoulou, 2012; Pyle et al., 2017). In addition, GOs can reveal the inferential relations among text elements (Graesser, Singer, & Trabasso, 1994) and facilitate students' skill in quickly locating specific information (Robinson & Skinner, 1996).

GOs vary in how strictly they map the text structure (Jiang & Grabe, 2007). Some GOs are previously established molds that more closely represent the discourse structure of the text, such as Venn diagrams for compare-and-contrast texts and timelines for chronological texts (Ocasio, 2006). These GOs are often instrumentally used to directly teach expository text structures (Alvermann & Boothby, 1986; Armbruster, Anderson, & Meyer, 1991). Typically, students read a text and fill in the empty slots in a partially completed GO, which afterward serves as input for a class discussion about text structures (Alvermann & Boothby, 1984; Boothby & Alvermann, 1984; Ermis, 2008; Moore, 1996; Van Steenbrugge, 2006). In some interventions, GO activities were complemented with writing down signaling words and main ideas (Broer et al., 2002), or students performed writing tasks based on information in their GOs (Moore, 1996; Raphael, Englert, & Kirschner, 1986).

Other GOs emphasize the hierarchical relations of textual information (e.g., between main ideas and supporting details), without focusing on a specific discourse structure. One example is mapping (Armbruster & Anderson, 1982; Berkowitz, 1986; Griffin et al., 1995), in which main ideas and their relevant relations are represented in a diagram.

For instance, Berkowitz (1986) taught students to write the title of a text in the middle of a sheet of paper, surrounded by boxes in which they noted one main idea per paragraph. Interventions were only included in our meta-analysis if the mapping involved representing the discourse structure or at least the hierarchical relations between ideas (e.g., main ideas vs. details).

In the context of narrative text structure instruction, story maps—schematic representations of the key information in narrative texts (Gardill & Jitendra, 1999)—are used. A story map can, for instance, include boxes for setting, goal, plot, and outcome, respectively (Tackett, Patberg & Dewitz, 1984). Story maps can be used postreading for summarization or during reading to help students monitor comprehension and/or highlight main events (Gardill & Jitendra, 1999; Gersten et al., 2001; Idol & Croll, 1987; Tackett et al., 1984). Story grammar can also help students formulate (Short & Ryan, 1984) or answer questions during reading that help in identifying the main constituents of the story (Gordon & Pearson, 1983). Because story maps also represent the structure of narrative texts, we included them in this meta-analysis.

So far, GO research has generated mixed results. Some studies have revealed positive results (e.g., Broer et al., 2002; Hoffman, 2010; Ulper & Akkok, 2010; Wijekumar et al., 2014; Wijekumar, Meyer, & Lei, 2012), but other interventions have been less effective (e.g., Alvermann & Boothby, 1984; Raphael et al., 1986) or showed that students need a great deal of instructional support to actually benefit from GOs (Griffin et al., 1995). One major issue concerns whether students benefit more from exposure to author-constructed or teacher-constructed organizers than from self-constructed organizers. Some researchers have stated that simple exposure to GOs may not be sufficient; rather, students may need extended instructional training and practice with GOs before they are able to recognize text structure and make use of this knowledge while reading (Jiang & Grabe, 2007). Various studies have shown, for instance, that the active involvement of students in constructing GOs, even when they are already partially complete, facilitates reading comprehension (e.g., Berkowitz, 1986; Spiegel & Barufaldi, 1994; Van Steenbrugge, 2006). However, Stull and Mayer (2007) argued that author-provided organizers are more effective, as student construction of GOs might create cognitive overload.

In this meta-analysis, we therefore examined the effects of structure-based visualizations on comprehension and analyzed whether active construction of structure-based GOs has a positive or negative additional impact on comprehension.

Structure-Based Summarization

To get the gist of a text, readers must overcome the limitations of working memory by ignoring extraneous or redundant

information and focusing specifically on macrolevel information, such as topic sentences (Bean & Steenwyk, 1984; Kintsch & van Dijk, 1978). This process of eliminating and reworking information can be promoted through summarization instruction (Armbruster et al., 1987; Bean & Steenwyk, 1984; Elledge, 2013; Frey, Fisher, & Hernandez, 2003; Taylor, 1985; Westby, Culatta, Lawrence, & Hall-Kenyon, 2010). Teaching summarization improves both the quality of written summaries and students' overall text comprehension (Duke, Pearson, Strachan, & Billman, 2011; Taylor & Beach, 1984). In their literature review, Miyatsu, Nguyen, and McDaniel (2018) stated that the defining characteristic of successful summarization training is the emphasis on main idea identification and text structure recognition. In fact, text structure can scaffold students' summarization skills, as it provides tools and heuristics to distinguish main ideas from unimportant information (Hogan, Bridges, Justice, & Cain, 2011; Meyer et al., 1980; E.A. Stevens, 2018; Taylor, 1985; Winograd, 1984). In addition, it helps students understand how these main ideas are organized, which helps them write coherent summaries (Miyatsu et al., 2018). Structure-based summarization strategies might also facilitate text recall, as the text structure might function as a mnemonic aid (Taylor, 1982).

In most elementary schools, teachers refrain from providing explicit instruction about specific summarization techniques (Beerwinkle et al., 2018; Elledge, 2013; McKeown, Beck, & Blake, 2009) or struggle themselves with the identification of main ideas (Kucan, Hapgood, & Palincsar, 2011; Turcotte et al., 2015). It is therefore not remarkable that many students struggle to identify main ideas (Baumann, 1983; Hare & Borchardt, 1984), seldom formulate summarizing topic sentences (Garner, 1987; Hare & Borchardt, 1984), and use deletion of propositional expression as their main summarization strategy (Winograd, 1984). Since the early 1980s, researchers have come up with various summarization techniques that are less intuitive, instead relying on the external or internal structure of texts.

The hierarchical summarization strategy (e.g., Taylor, 1982, 1985) consists of skimming the external organizational text structure (i.e., the headings and subheadings) first and then preparing a skeletal outline based on the headings. Next, students write one main idea per section in their outline. However, the strategy might be limited to texts with an unambiguous heading/subheading format (Armbruster et al., 1987). The rule-based summarization strategy (Brown, Day, & Jones, 1983; McNeil & Donant, 1982) relies more heavily on the internal structure of paragraphs (i.e., identifying topic sentences). This strategy provides students with a set of six summarization rules, based on the work of Kintsch and van Dijk (1978), such as delete redundant information and invent a topic sentence. These rules help students first

eliminate information and then rework the remaining bits into a coherent summary (Brown et al., 1983; McNeil & Donant, 1982). For instance, students are taught to highlight topic sentences, circle words that must be replaced by superordinate concepts, and cross out trivia (Braxton, 2009). Rule-based summarization is often combined with text structure recognition. For instance, students learn how signal words and structure-specific questions can be used to identify main ideas (e.g., Elledge, 2013).

Not only the internal structure of the paragraph but also that of the whole text can function as a framework for summarization (Armbruster et al., 1987; de Jou & Sperb, 2009; Ocasio, 2006; E.A. Stevens, 2018; Vidal-Abarca, 1990). For instance, in the studies by Armbruster et al. (1987) and E.A. Stevens (2018), students were taught about the characteristics of the problem-and-solution text structure. Then, students received a specific problem-and-solution frame in which they could summarize the main point of a text, as well as a list of structure-specific guidelines for their summary (e.g., "Sentence 1 - Tells who had a problem and what the problem is"; Armbruster et al., 1987, p. 337). Vidal-Abarca (1990) and de Jou and Sperb (2009) explained various text structures and modeled where and how to find the main idea in these texts.

In this meta-analysis, we examined the effects of structure-based summarization techniques. In particular, we analyzed whether the rule-based summarization approach, with its emphasis on internal text structure, has an additional impact on text comprehension.

Instructional Variables

Teaching students about text structures can be done in many different ways. Previous meta-analyses have shown that the implementer plays a crucial role. Researcher-taught interventions are often more effective than teacher-taught interventions (e.g., Dignath & Büttner, 2008), such as in text structure instruction (Pyle et al., 2017).

Over the past decade, more attention has been paid to the ecological component of text structure instruction by investigating how teachers explain text structure and other evidence-based strategies in their classroom (Beerwinkle et al., 2018; Wijekumar et al., 2019) and how teachers can be trained to teach text structures (Reutzel, Jones, Clark, & Kumar, 2016). Meyer and Ray (2011) emphasized that teachers should have access to adequate instructional materials for modeling and direct instruction, such as by providing them with "intelligent tutors or scripted lessons" (p. 138). Williams (2018) also made a plea for second-generation text structure research that goes beyond developing excellent instructional materials and focuses more on the context in which the instruction occurs. In addition, the previous meta-analyses stated that future research should examine the mediating role of instructional features, as

both the implementer and the type of instructional activities might affect the effectiveness of text structure interventions (Hebert et al., 2016; Pyle et al., 2017).

The gradual release of responsibility model (Fisher & Frey, 2013; Pearson & Gallagher, 1983) gives a useful framework for describing and comparing current teaching practices. This model suggests sequencing various instructional activities such that the responsibility for the learning process relies mainly on the teacher first (e.g., direct instruction, modeling) and is then gradually transferred to the student with decreasing levels of scaffolding (e.g., guided practice, collaborative activities, individual activities). This also reflects the idea that reading comprehension lessons should follow a pattern of stepwise phasing out the teacher while phasing in the students. They gradually take over the lead from the teacher by applying comprehension strategies first in small groups, then in pairs, and finally individually (Nolte & Singer, 1985; Singer & Donlan, 1980). Many studies on reading instruction have emphasized the importance of direct instruction of cognitive and metacognitive strategies for reading (e.g., Rosenshine & Meister, 1994; R.J. Stevens, Slavin, & Farnish, 1991).

Research has suggested that modeling also plays a pivotal role in increasing reading comprehension, especially when the demonstration of the model includes conditional knowledge: why the model is doing something, as well as metacognitive and motivational aspects (Kostons, Donker, & Opdenakker, 2014). Modeling can raise students' self-efficacy to carry out tasks on their own (van Gog & Rummel, 2010), which may be particularly beneficial to students with low reading self-efficacy. In various studies, positive effects have been found for explicit instruction combined with teacher modeling (Duffy, 2002).

For reading and many other areas of instruction, collaboration has been a successful way to enhance learning. If well implemented, it can improve students' time on task (E.G. Cohen & Benton, 1988), raise academic performance (Slavin, 1987), and increase the quantity and quality of student interactions (Fuchs, Fuchs, Mathes, & Simmons, 1997; Garibaldi, 1979; Vaughn, Klingner, & Bryant, 2001), among other benefits.

Text structure interventions in the upper elementary grades have shown variability in terms of their instructional approach. Most interventions are composed of direct instruction (with or without modeling) followed by individual practice (e.g., Armbruster et al., 1987; Berkowitz, 1986; Bowman & Gambrell, 1981; Elledge, 2013; Gordon & Pearson, 1983; Griffin et al., 1995), whereas in other studies, students only worked in groups (e.g., Darch, Carnine, & Kameenui, 1986) or both in groups and individually (e.g., Fitzgerald & Spiegel, 1983; Moore, 1996; Scott, 2011; Tackett et al., 1984; Van Steenbrugge, 2006). Only a handful of studies displayed the full gradual release of responsibility with explicit instruction, modeling, and collaborative and individual work (e.g., Braxton, 2009; Ermis, 2008; Gentry,

2006). Researchers have focused on the development of web-based text structure teaching, such as intelligent tutoring of the structure strategy (e.g., Meyer et al., 2010; Wijekumar et al., 2012, 2014; Wijekumar, Meyer, & Lei, 2013), and on the effectiveness of specific instructional features, such as the presence or absence of tutors (Meyer et al., 2002) or instruction tailored to students' individual needs (Meyer et al., 2010).

Except for analyzing the effects of online tutors versus teachers (Pyle et al., 2017), prior meta-analyses did not analyze the effects of instructional components on the effectiveness of text structure instruction. Therefore, in our meta-analysis, we examined whether it matters if teacher and student activities follow a pattern of gradual release of responsibility.

Method

Inclusion and Exclusion Criteria

We developed a set of inclusion and exclusion criteria to guide our iterative search and selection procedure. Studies were deemed eligible for inclusion in the meta-analysis if the following five criteria were met:

1. The study was published in English, German, French, Dutch, or Spanish between 1974 and 2018 and was available online or could be retrieved directly from the author.
2. The study focused on students in general primary education in grades 4–6. Participants could be students with mild reading difficulties but not students with severe learning or reading difficulties, students with hearing problems, and/or second-language learners.
3. The treatment group was taught to recognize informational text structures (e.g., description, compare-and-contrast, problem-and-solution, cause-and-effect, sequence; see Meyer, 1975) or story structure. Interventions in which text structure was explicitly used as part of a summarization or visualization technique were also included.
4. The treatment group was compared with a business-as-usual control group or a control group receiving an alternative instruction. There was no restriction on sample size or sampling procedures as long as the procedure was well documented.
5. The study included at least one posttest focused on the comprehension of texts. The posttest could consist of one or several (non)standardized comprehension question tests, cued or free recalls, GO completion, and/or summarization tasks and text structure knowledge tests.

Studies were excluded if the researcher(s) did not provide the statistics necessary to calculate a weighted effect size or if results were summarized over multiple age groups, such that it was impossible to calculate effect sizes for grades 4–6 exclusively, even after contacting those authors. Correlational and qualitative studies were excluded, as well as studies with a within-subjects design or a multiple-baseline design (e.g., Haria & Midgette, 2014).

Search and Selection Procedure

We used a four-step process to conduct a comprehensive search for text structure intervention studies. First, we located studies by using the electronic databases of ERIC, Web of Science, PsycINFO, and Google Scholar. Second, we searched various databases of theses and dissertations (e.g., PQDT Open, EThOS, OpenThesis) to locate unpublished studies. Third, we used a series of Dutch, French, German, and Spanish search engines to identify relevant studies in languages other than English. Fourth, we conducted a cited reference search of previous reviews and meta-analyses (Hebert et al., 2016; Jiang & Grabe, 2007; Meyer & Ray, 2011; Pyle et al., 2017) and checked reference lists of the studies that were judged eligible.

In the literature search, we used the key terms *text structure* or *top-level structure* and *reading (comprehension)* combined with keywords specifying the age group (i.e., *primary, elementary, fourth/fifth/sixth grade*). In a second search, we also included search terms about the type of text structure (i.e., *description, enumeration, classification, cause effect, compare contrast, sequence, chronology, problem solution, story (grammar), narrative*) and about the potential ingredients in text structure interventions (i.e., *topic sentence, signaling words, cue words, graphic organizer, schematizing, main idea, outline, summary*) to maximize the number of articles located. This search yielded approximately 2,900 results.

After removal of the many duplicates from the list and a first title screening, we exported 408 abstracts for closer examination. We read abstracts and Method sections to determine whether these studies qualified for inclusion. We removed 355 articles because participants did not match the required age and aptitude profile ($n = 78$) or because studies did not meet the criteria for text structure instruction ($n = 208$; e.g., reciprocal teaching interventions or semantic mapping studies without a focus on text structure), were focused on text manipulations ($n = 45$), did not measure text comprehension ($n = 20$), or were published in another language ($n = 4$), which left us with 52 eligible studies. Eight studies could not be located or provided insufficient statistics for effect size calculation, such that the final set consisted of 44 studies published between 1982 and 2018.

Study Feature Coding

We coded studies on publication type, participants (number and grade), research design (quasi-experimental vs. experimental), type of outcome measure (i.e., comprehension questions, free recall tasks, summarization tasks, text structure knowledge tests), test type (standardized vs. nonstandardized), measurement timepoint (immediate vs. delayed), and reference group (business as usual vs. alternative intervention as control). We also coded text features, such as genre (narrative and/or informational) and the number and exact type of structures (e.g., narrative, description, cause-and-effect, compare-and-contrast, problem-and-solution, sequence). In addition, we coded various content-related and instructional features of both the treatment and control conditions (see Table 1).

We described the instructional content of the interventions by using three higher level descriptors that formed the basis for further analysis. These descriptors were not mutually exclusive: One intervention, for instance, could include both training in text structure recognition and structure visualization. To refine these categories, we also added variables describing whether there was an additional focus on paragraph-level structure, active construction of GOs, and/or instruction in the rule-based summarization technique.

We defined two instructional variables: teacher activities and student activities. Instruction activities could be primarily focused on telling (i.e., explicit instruction) or showing (i.e., modeling). Student activities were either individual or collaborative. The analysis of these instructional components was based on the description of the intervention and procedures in the article or on an examination of the examples of materials that were provided (e.g., screenshots, examples, teaching materials in appendixes). If the text mentioned that teachers used think-aloud protocols or demonstrated/ modeled the strategy, this was coded as modeling. Similarly, if the authors mentioned working in pairs or small-group activities, we

scored this as a collaborative activity. If nothing was mentioned or could be derived logically, this was coded as absence of the instructional feature at hand.

Apart from the distinction between standardized and nonstandardized tests, we also distinguished among four types of outcome measures: comprehension questions, free recall tasks, summarization tasks, and text structure knowledge tests. Outcome measures coded as comprehension questions were all tests that involved multiple-choice items and short-answer questions with the text present (e.g., literal comprehension, referential comprehension, interpretation questions). Recall tasks included tasks in which students were typically asked to read a text and then perform a memory task without the text present. This could be writing or telling everything that they remembered or only the most important information, or cued recall tasks in which they answered factual questions about text content (e.g., where did the story take place?). Summarization tasks included all tasks in which students had to summarize part of a text with the text present, such as student-generated written summaries, GO completion, and tasks focused on highlighting the main ideas. Text structure knowledge tasks included all tests focused on knowledge about specific text structures, such as tasks in which students had to identify the structure of a text segment (e.g., E.A. Stevens, 2018), match segments to a text with a similar structure (e.g., Broer et al., 2002), or select the right structure-specific signaling words in cloze tasks (e.g., Wijekumar et al., 2013).

All studies were coded by the first author. A research assistant was trained in the coding procedure and coded a random sample of four studies (10% of the total sample), with 94% total interobserver agreement and Cohen's kappa values ranging from .83 (instructional approach) to .95 (methodological descriptors). Table 2 provides an overview of all studies and shows immediate and delayed effect sizes per outcome measure.

TABLE 1
Description of Intervention Content and Instructional Components

Descriptor	Description
Structure recognition and/or focus on the paragraph level	Students are explicitly taught about the internal structure of texts or paragraphs and practice with recognition. The focus on the paragraph level might entail instruction about topic sentences and supporting and concluding sentences.
Structure visualization and/or active construction	Students receive instruction on schematic representations of text structure and content. They study and/or fill out or actively create maps and graphic organizers.
Structure-based or rule-based summarization	Students learn how to summarize a text on the basis of headings and other hierarchical outlining principles or learn techniques for paraphrasing main ideas or condensing text by strictly following a set of rules.
Teacher-led instruction ± modeling	The teacher provides explicit instruction and/or demonstrates structure recognition, structure visualization, or structure-based summarization techniques by thinking aloud in front of the class.
Student activities ± individual practice	Students have opportunities to practice their skills, collaboratively and/or individually.

TABLE 2
Descriptive and Effect Size Information per Study and per Outcome Measure

Study	Publication type ^a	Grade(s)	N ^b	Intervention(s) vs. control	Genre(s) ^c	Number of effect sizes	Effect sizes (Hedges' g)			
							Comprehension questions	Recall	Summarization	Text structure knowledge
<i>Text structure recognition (n = 20)</i>										
Bohaty (2015)	D	4 and 5	45 L	Text structure recognition and focus on cue words vs. business as usual	CC, CE, D, PS, S	3	-0.17 ^d	0.13		0.99
Carriedo and Alonso-Tapia (1996)	J	6	211	TSI with a strong focus on main idea identification vs. traditional questioning	CC, D, N, P, PS	2			0.58	-0.24
Fitzgerald and Spiegel (1983)	J	4	19	Story grammar vs. vocabulary and dictionary instruction	N	6	-0.13	-0.02		
Gentry (2006)	D	4	30	Recognizing structures and with text annotation vs. vocabulary instruction	CC, CE, D, PS, S	2	0.11 vs. -0.08 ^d			
Gordon and Pearson (1983)	R	5	42 H	Story grammar vs. inference awareness	N	6	0.29 vs. 0.43 ^d	0.69		
McDermott (1991)	D	4	66	Recognizing structures vs. business as usual	D, PS	2		0.66		
Meyer et al. (2002)	J	5	60	Web-based structure strategy (± tutoring) vs. business as usual	CC, CE, D, PS, S	16	0.19 vs. 0.61 ^d	0.41 vs. 0.05 ^d	0.30 vs. 0.15 ^d	0.50 vs. 0.43 ^d
Meyer, Wijekumar, and Lin (2011)	J	5	131	Web-based structure strategy (individualized instruction) vs. web-based structure strategy (standardized)	CC, PS	12	0.42	0.08	-0.07	0.45
Ponce, López, and Mayer (2012)	J	4	1,041	Recognizing structures (online) vs. business as usual	CC, CE, P, S	1	0.71			
Raphael, Englert, and Kirschner (1986)	R	5 and 6	159	TSI (writing ± social context) vs. business as usual	CC, D, N, PS	2	0.06		0.38	
Raphael and Kirschner (1985)	R	5 and 6	45	TSI with writing vs. business as usual with writing	CC	2		1.22	0.82	

(continued)

TABLE 2
Descriptive and Effect Size Information per Study and per Outcome Measure (continued)

Study	Publication type ^a	Grade(s)	N ^b	Intervention(s) vs. control	Genre(s) ^c	Number of effect sizes	Effect sizes (Hedges' g)			
							Comprehension questions	Recall	Summarization	Text structure knowledge
Scott (2011)	D	6	205	TSI vs. content-focused instruction (social studies)	D, S	4	-0.27	0.95		
Short and Ryan (1984)	J	4	56 L	Story grammar vs. attribution training	N	2	0.62	0.66		
Spires, Gallini, and Riggsbee (1992)	J	4	74	TSI vs. previewing structure vs. traditional questioning	CC, PS	16	0.40 vs. 0.08 ^d	0.04 vs. -0.35 ^d		
Tackett, Patberg, and Dewitz (1984)	J	6	45 L	Story grammar vs. exposure to well-formed stories vs. business as usual	N	2	0.54			
Troyer (1994)	P	4-6	173	TSI and mental models vs. TSI and graphic organizers vs. TSI vs. business as usual	CC, D	14	0.37 vs. 0.20 ^d		0.40	
Walker (1995)	J	5	105	TSI vs. traditional questioning	CC, CE, D, PS, S	2	0.09 vs. -0.15 ^d			
Wijekumar, Meyer, and Lei (2012)	J	4	2,600	Web-based structure strategy vs. business as usual	CC, CE, D, PS, S	7	0.10	0.49	0.20	
Wijekumar, Meyer, and Lei (2013)	J	4 and 5	443	Web-based structure strategy vs. business as usual	CC, CE, D, PS, S	8	0.19	0.74	0.39	
Wijekumar et al. (2014)	J	5	2,645	Web-based structure strategy vs. business as usual	CC, CE, D, PS, S	7	0.20	0.53	0.29	

(continued)

TABLE 2
Descriptive and Effect Size Information per Study and per Outcome Measure (continued)

Study	Publication type ^a	Grade(s)	N ^b	Intervention(s) vs. control	Genre(s) ^c	Number of effect sizes	Effect sizes (Hedges' g)			
							Comprehension questions	Recall	Summarization	Text structure knowledge
<i>Text structure visualization (n = 13)</i>										
Alvermann and Boothby (1986)	J	4	24	GO training (14 weeks) vs. GO training (7 weeks) with some TSI vs. reading recitation	D	9	0.23	0.43		
Armbruster, Anderson, and Meyer (1991)	J	4 and 5	365	GO training vs. business as usual	E	2	0.11			
Berkowitz (1986)	J	6	99	GO construction vs. GO study vs. questioning vs. rereading	E	40	0.08 vs. -0.05 ^d	0.46 vs. 0.05 ^d		
Boothby and Alvermann (1984)	J	4	38	GO training vs. no specific instruction	CE	5	0.99	0.98 vs. 0.17 ^d		
Bowman and Gambrell (1981)	P	6	100	Story maps and questioning vs. traditional questioning	N	4	0.46 vs. 0.08 ^d	0.19 vs. 0.32 ^d		
Broer, Aarnoutse, Kieviet, and van Leeuwe (2002)	J	6	354	Recognizing structures with emphasis on schematics vs. traditional questioning	CE, D	7	0.15	0.55 vs. 0.18 ^d		0.48
Darch, Carmine, and Kameenui (1986)	J	6	84	GO training (group) vs. directed reading	E	4	0.71			
Ermis (2008)	C	4 and 5	50	GO training (individual) vs. directed reading	CE, D	2	0.67			
Griffin, Malone, and Kameenui (1995)	J	5	99	GO training with some TSI vs. traditional questioning	E	24	0.35	0.04		
Hoffman (2010)	D	5	153	Explicit vs implicit main idea training			-0.25 ^d	-0.42 ^d		
McLaughlin (1990)	D	5	68 L	GO training and metacognitive monitoring vs. no instruction vs. GO training	CC	4	0.33			
Moore (1996)	P	6	76	GO training with TSI vs. other cognitive strategies	CC	3	0.10	0.04		0.21
Van Steenbrugge (2006)	M	5 and 6	103	GO training and TSI vs. traditional questioning	CC, CE	2	0.72 vs. 0.40 ^d			
				GO construction vs. GO study vs. business as usual	CE, D, N, S	3	0.05			

(continued)

TABLE 2
Descriptive and Effect Size Information per Study and per Outcome Measure (continued)

Study	Publication type ^a	Grade(s)	N ^b	Intervention(s) vs. control	Genre(s) ^c	Number of effect sizes	Effect sizes (Hedges' g)			
							Comprehension questions	Recall	Summarization	Text structure knowledge
<i>Structure-based summarization (n = 11)</i>										
Armbruster, Anderson, and Ostertag (1987)	J	5	82	Rule-based summarization and TSI vs. traditional questioning	PS	2	0.64		1.66	
Bean and Steenwyk (1984)	J	6	60	Rule-based summarization vs. intuitive summarization vs. no specific instruction		4	0.80		0.74	
Braxton (2009)	D	4 and 5	64	Rule-based summarization vs. intuitive summarization	CC, CE, D, PS, S	4	-0.23		1.82	
de Jou and Sperb (2009)	J	5	86	Main idea identification training with some TSI and metacognitive strategies vs. business as usual	E, N	4	0.35	0.66		
Elledge (2013)	D	4-6	126	Rule-based summarization and TSI vs. business as usual	CE, D	6			0.10	
Ocasio (2006)	D	5	63	Structure-based summarization and TSI vs. reciprocal teaching	CC, CE, PS, S	1			2.36	
E.A. Stevens (2018)	D	4 and 5	61 L	Main idea identification and TSI	CC, D, PS	3	-0.14	0.70		0.58
Taylor (1982)	J	5	48	Hierarchical summarization vs. traditional questioning	E	2	-0.07			0.91
Taylor (1985)	J	6	93	Hierarchical summarization ± TSI with writing vs. traditional questioning and writing	CC	18	0.27 vs. -0.47 ^e	0.25 vs. 0.51 ^d	0.39 vs. -0.01 ^d	
Ulper and Akkok (2010)	C	6	13 L and 15 H	Structure-based summarization and some TSI vs. structure-based summarization vs. business as usual	S	2			0.90	
Vidal-Albarca (1990)	J	5	51	Main idea identification and some TSI vs. business as usual		3			1.06	1.04

Note. GO = graphic organizer; TSI = text structure instruction.

^aC = book chapter; D = doctoral dissertation; J = Journal article; M = master's thesis; P = paper; R = report. ^bH = high-achieving students; L = low-achieving students. ^cCC = compare-and-contrast; CE = cause-and-effect; D = descriptive; E = expository (no further specification); N = narrative; P = persuasive; PS = problem-and-solution; S = sequence. ^dA delayed effect (i.e., immediate posttest-delayed posttest difference).

Effect Size Calculation and Statistical Analyses

For calculation of effect sizes, we used Hedges' g , which is almost similar to Cohen's d (J. Cohen, 1988; Fritz, Morris, & Richler, 2012) but provides an unbiased estimate of effect sizes for the few studies with smaller sample sizes ($n < 20$) that were included in this meta-analysis (Borenstein, Hedges, Higgins, & Rothstein, 2009; Cumming, 2012). Although effect sizes of approximately 0.20 are generally classified as small, we interpret them as meaningful because they were obtained in educational contexts, where even effect sizes that are generally classified as small are of interest (see Durlak, 2009; Hedges & Hedberg, 2007). Moreover, the effect sizes that we report in this meta-analysis should be interpreted as additive effects because they represent what students in a text structure condition gained on top of what students learned in an alternative intervention or the regular reading curriculum.

For each study, we calculated the standardized mean differences (Hedges' g) for immediate and delayed effects separately. For the calculation of immediate effects, we subtracted the mean difference in performance of the control group (immediate posttest – pretest) from the mean difference in performance of the treatment group (immediate posttest – pretest), divided by the pooled standard deviation of the two groups. As we were interested in long-term differences between experimental and control groups, we also calculated delayed effect sizes. These are indicative of the effects of text structure instruction that remain, over and above what students learned in business as usual. This calculation was based on comparing the mean performance of the control group versus the treatment group on delayed posttests, divided by the pooled standard deviation of the two groups. A delayed effect of 0 would mean that there were no lasting differences between both groups, whereas a delayed effect of +1 would indicate that the experimental group outperformed the control group by one standard deviation at delayed posttests, either because the experimental group made more progress than the control group or because the control group performed worse while the experimental group maintained their acquired skill with respect to the immediate posttest.

For the six studies that did not report the exact means and standard deviations, we calculated effect sizes based on analyses of variance. Because the sample sizes of these six studies were quite large ($n > 42$), the calculated value of Cohen's d was practically identical to the Hedges' g value (Borenstein et al., 2009; Fritz et al., 2012).

When multiple outcome measures were used, we calculated separate effect sizes per type of outcome measure (i.e., comprehension questions, recall, summarization, text structure knowledge) so we could show the impact of the intervention variables on each measure of text

comprehension separately. Also, we calculated separate effect sizes when multiple text structure interventions were compared within studies. For instance, Ulper and Akkok (2010) investigated the effectiveness of structure-based summarization strategies and the effectiveness of this approach in combination with training in text structure recognition. In these instances, we calculated separate effect sizes per condition, even though these are presented as averaged effect sizes in Table 2.

Because the handling of multiple effect sizes from one study led to statistical dependencies in the data, we aggregated data sets per outcome measure. By taking this approach, we calculated no more than two effect sizes per study: one delayed and one immediate. In the analyses, we allocated more weight to studies with larger sample sizes (Borenstein et al., 2009). We ran all mixed-effect model analyses per outcome measure, so we actually conducted four parallel meta-analyses, in which we subsequently added the aforementioned methodological, content-related, and instructional variables as moderators. After constructing full factorial models, we simplified the models with only the relevant parameters so we could estimate the additive effect of all parameters, without running the risk of overfitting (see the tables in the Appendix for more details).

Due to differences in types of participants, and methodological, content-related, and instructional characteristics, we could not make the assumption of one common effect size. Therefore, we used random effects models, which assume not one true effect size but an effect size distribution. This made it possible for us to generalize to populations beyond the included studies (Borenstein et al., 2009). We also examined the within-class goodness of fit by conducting homogeneity tests (Cooper, 1998) to check whether the variability in effect sizes was so large that moderator analyses were needed. We tested the differences in fit of subsequent (nested) random effects models by means of log likelihood ratio tests. All effect size calculations and moderator analyses were conducted in R using the metafor package (version 3.3.3; Viechtbauer, 2010). A full overview of our stepwise model fitting can be found in the Appendix.

Results

For each outcome measure, we used a random-effects model to assess the overall average effect size. The overall effect size for comprehension questions was $g = 0.14$, 95% confidence interval (CI) [0.03, 0.25]; for recall, $g = 0.30$, 95% CI [0.19, 0.41]; for summarization, $g = 0.43$, 95% CI [0.24, 0.61]; and for text structure knowledge, $g = 0.34$, 95% CI [0.28, 0.41]. We had to be careful with the interpretation of these mean effect sizes, as for all outcome

measures, there was statistically significant heterogeneity in effect sizes, $Q_{\text{Questions}}(88) = 310.37, p < .001$; $Q_{\text{Recall}}(84) = 248.38, p < .001$; $Q_{\text{Summarization}}(51) = 328.47, p < .001$; $Q_{\text{Knowledge}}(26) = 79.87, p < .001$. Of course, this heterogeneity can partly be attributed to the fact that the overall effect sizes still include both immediate and delayed effects and various types of research designs.

For comprehension questions and text structure knowledge, we were unable to show a publication bias, $\Delta\chi^2(1) = 1.32, p = .25$, and $\Delta\chi^2(1) = 0.23, p = .63$, respectively. However, for recall and summarization, we found evidence for a reversed publication bias, $\Delta\chi^2(1) = 5.00, p = .03$, and $\Delta\chi^2(1) = 3.95, p = .047$, respectively. That is, effect sizes for recall and summarization reported in peer-reviewed scientific journals were systematically smaller than in other publications, including book chapters and nonpublished dissertations.

Methodological Variables

Control Condition

In some studies ($n = 18$), the authors explicitly mentioned that the text structure intervention was compared with an alternative instruction instead of to a business-as-usual control group. For instance, text structure instruction was compared with vocabulary instruction (e.g., Fitzgerald & Spiegel, 1983; Gentry, 2006), instruction in cognitive reading strategies such as making predictions and inferences and activating prior knowledge (e.g., Gordon & Pearson, 1983; McLaughlin, 1990; Ocasio, 2006), or more intuitive summarization strategies (Bean & Steenwyk, 1984; Braxton, 2009). However, in most studies, the control groups continued to follow their usual reading curriculum during the intervention. Business as usual typically involved a traditional approach to reading instruction, with students answering questions about the text and the teacher leading class discussions focused on text content, although business as usual also contained elements prevailing in the aforementioned alternative instruction programs (e.g., activating prior knowledge, explaining vocabulary). Both types of control groups occurred with all types of outcome measures.

Stepwise model fitting showed that effect sizes were not systematically different when the intervention was compared with an alternative program instead of business as usual: comprehension questions, $\Delta\chi^2(1) = 1.18, p = .28$; recall, $\Delta\chi^2(1) = 0.44, p = .51$; summarization, $\Delta\chi^2(1) = 2.16, p = .14$; and text structure knowledge, $\Delta\chi^2(1) = 1.14, p = .29$. For example, the estimated effect size of text structure instruction on recall was $g = 0.32$ (standard error [SE] = 0.06) when compared with an alternative instruction, and $g = 0.24$ (SE = 0.12) when compared with business as usual, which was not statistically significant. For the remaining analyses, we therefore did not distinguish between the two types of control conditions.

Standardized and Nonstandardized Measures

We used both standardized and nonstandardized measures to evaluate the effects of text structure instruction. For comprehension questions, approximately 22% of the effect sizes were based on standardized measures. For recall, summarization, and text structure knowledge tests, only nonstandardized measures were used. We could not demonstrate an effect of standardization, $\Delta g = -0.2$, $SE = 0.11$; $\Delta\chi^2(1) = 0.04, p = .84$. Differences in effect sizes due to text structure instruction were similar when measured with standardized and nonstandardized comprehension questions.

Immediate and Delayed Effects

Not all studies provided data on the maintenance of effects. Delayed posttests were administered in approximately one third of the studies ($n = 16$). Most of these concerned comprehension questions ($n = 12$) and recall ($n = 6$) and took place two or three weeks after completing the intervention (63%). The immediate effect sizes were above 0.20 for all outcome measures and were therefore meaningful: Students who received text structure instruction outperformed the control group on comprehension questions ($g = 0.25, SE = 0.07$), recall ($g = 0.38, SE = 0.06$), summarization ($g = 0.58, SE = 0.09$), and text structure knowledge ($g = 0.34, SE = 0.03$). However, there was a statistically significant decrease in delayed effects (see Table 3): comprehension questions, $\Delta\chi^2(1) = 6.89, p < .001$; recall, $\Delta\chi^2(1) = 4.83, p = .03$; and summarization, $\Delta\chi^2(1) = 9.73, p = .002$. For text structure knowledge, no delayed effects could be calculated because of a limited number of observations.

In fact, for each of the outcome measures, the differences between groups with or without text structure instruction completely disappeared at the delayed posttests, as none of the estimated effect sizes reached significance: comprehension questions, $g = -0.05, SE = 0.11$; recall, $g = 0.13, SE = 0.11$; and summarization, $g = -0.06, SE = 0.20$. In other words, although the students in the text structure condition outperformed the controls on each outcome measure at the immediate posttest, this difference between conditions was not maintained at the delayed posttest, where students' performance in both conditions was similar. In most cases, the performance of the experimental groups showed a stronger decrease after the immediate posttest than that of the control groups, whose performance remained rather stable or showed a small decrease.

The remaining effect sizes were still heterogeneous after adding the previously mentioned variables, $Q_{\text{Questions}}(87) = 300.18, p < .001$; $Q_{\text{Recall}}(83) = 245.08, p < .001$; $Q_{\text{Summarization}}(50) = 276.73, p < .001$; and $Q_{\text{Knowledge}}(25) = 79.26, p < .001$. Thus, we decided to

TABLE 3
Immediate and Delayed Effect Sizes per Outcome Measure

Outcome measure	Immediate effect size		Delayed effect size		Chi-square test
	Hedges' <i>g</i>	Standard error	Hedges' <i>g</i>	Standard error	
Comprehension questions	0.25	0.07	-0.05	0.11	$\Delta\chi^2(1) = 6.89, p < .001$
Recall	0.38	0.06	0.13	0.11	$\Delta\chi^2(1) = 4.83, p = .03$
Summarization	0.58	0.09	-0.06	0.20	$\Delta\chi^2(1) = 9.73, p = .002$
Text structure knowledge	0.34	0.03			

Note. Delayed effect sizes are the difference between delayed and immediate posttest scores. Due to power insufficiencies ($n = 3$), no delayed effect sizes were calculated for text structure knowledge.

conduct additional moderator analyses focusing on content-related and instructional variables.

Text Variables

To determine whether the type and number of text structures affected the effect sizes, we successively added genre (informational or narrative) and the number of different text structures taught to the four models.

Genre

Most studies discussed the effects of teaching informational text structures ($n = 34$). In some studies, students also received instruction in narrative story structure ($n = 5$) or in narrative story structure only ($n = 5$). Interventions on informational text structures included all types of outcome measures, whereas interventions involving narrative story structure mainly used comprehension questions and recall tasks. We could not demonstrate an effect of genre on these outcome measures: comprehension questions, $\Delta\chi^2(1) = 0.12, p = .73$; and recall, $\Delta\chi^2(1) = 0.11, p = .74$. There was also no interaction effect of genre on delayed posttest performance: comprehension questions, $\Delta\chi^2(1) = 0.89, p = .34$; and recall, $\Delta\chi^2(1) = 0.11, p = .74$. Interventions focusing on informational text structures only versus studies (also) including narrative texts were comparable in terms of their effects on text comprehension. Because the number of studies focusing on narrative story structure was limited, we could not examine genre effects on summarization and text structure knowledge.

Number of Text Structures

Some intervention studies focused on only one text structure ($n = 20$), whereas other studies taught up to five structures ($n = 9$). Overall, the description and compare-and-contrast text structures were taught most frequently. More text structures were taught in interventions with a text structure knowledge test (mean [M] = 3.81, standard deviation [SD] = 1.47) or a summarization task ($M = 2.37, SD = 1.44$) as the outcome

measure than in interventions with comprehension questions ($M = 1.53, SD = 1.07$) or recall tasks ($M = 1.64, SD = 1.35$). There was a small negative effect of the number of different text structures on text structure knowledge, $\Delta g = -0.06, \Delta\chi^2(1) = 4.64, p = .03$; but not on comprehension questions, $\Delta\chi^2(1) = 0.54, p = .46$; recall, $\Delta\chi^2(1) = 0.003, p = .96$; or summarization, $\Delta\chi^2(1) = 0.21, p = .65$. The more different text structures were taught during an intervention, the lower the scores on the text structure knowledge test were (i.e., $\Delta g = -0.06$ times the number of different text structures taught; we could not show a curvilinear effect). The number of text structures taught did not matter when students made a summary, answered comprehension questions, or carried out a recall task.

We also checked for an interaction effect between the number of text structures taught and the measurement timepoint to see if the number of text structures mattered for students' performance on delayed posttests. This interaction effect was not found for any of the outcome measures: comprehension questions, $\Delta\chi^2(2) = 3.44, p = .18$; recall, $\Delta\chi^2(2) = 0.02, p = .99$; summarization, $\Delta\chi^2(2) = 0.21, p = .90$; and text structure knowledge, $\Delta\chi^2(1) = 4.57, p = .03$. The number of different text structures taught did not affect maintenance effects.

Content Features and Instructional Components

We analyzed whether the effects of text structure instruction were affected by the content features and instructional components listed in Table 1. Because of a limited number of observations, we did not perform this moderator analysis for text structure knowledge. All final models were an improvement relative to the models without moderating content-related and instructional variables: comprehension questions, $\Delta\chi^2(5) = 14.13, p = .015$; recall, $\Delta\chi^2(5) = 15.49, p = .008$; and summarization, $\Delta\chi^2(3) = 20.81, p = .001$. The four final models explain 25–27% of the variance in effect sizes. Next, we discuss the outcomes of these moderator analyses.

Content Features

Table 4 shows the estimated immediate effects per outcome measure and the estimated additional effects of various content features and instructional components. The parameter estimates from the final models (see the Appendix) show that not all features contributed evenly to the effects on the different outcome measures. Training in text structure recognition had a statistically significant effect on answering comprehension questions ($g = 0.98$, $SE = 0.30$, $p = .001$) and recall ($g = 1.03$, $SE = 0.39$, $p = .009$). For recall, instruction on paragraph-level structure also mattered ($\Delta g = 0.57$, $SE = 0.29$, $p = .03$), but this was not the case for comprehension questions. For summarization skills, it was specifically a focus on paragraph-level structure that mattered ($\Delta g = 0.91$, $SE = 0.22$, $p < .001$), whereas training in only top-level text structure recognition did not improve students' summarization skills ($\Delta g = 0.22$, $SE = 0.40$, $p = .58$). Apparently, when it comes to summarizing, students benefit most from text structure instruction that also focuses on the internal structure of paragraphs ($g = 0.91 + 0.22 = 1.13$).

General attention to structure visualizations had no demonstrable impact on students' performance on text comprehension questions unless intervention programs emphasized the actual construction of GOs and story maps ($\Delta g = 0.39$, $SE = 0.15$, $p = .009$). For recall, simple exposure to GOs had a negative effect ($\Delta g = -0.44$, $SE = 0.18$, $p = .02$), whereas active construction had a positive effect ($\Delta g = 0.51$, $SE = 0.17$, $p = .002$). This shows that when interventions asked students to actively create or fill out maps or GOs in addition to practicing text structure recognition, this had an effect on comprehension questions ($g = 0.64$) and recall ($g = 1.03$).

Structure-based summarization training in general had no demonstrable additional effect on comprehension

question answering, recall, or summarization skills over and above training in text structure recognition. However, specific training in the rule-based summarization technique positively affected summarization skills ($\Delta g = 0.64$, $SE = 0.21$, $p = .005$) and recall ($\Delta g = 0.34$, $SE = 0.12$, $p = .004$) but had no statistically significant impact on comprehension questions. When interventions trained students to apply a fixed set of structure-based rules to summarize text, this resulted in net immediate effects of $g = 0.73$ on summarization and $g = 1.12$ on recall.

None of the content-related features had a demonstrable impact on delayed posttest performance. As we reported earlier, the differences between students receiving text structure instruction and students in the control groups that were visible immediately after finishing the intervention program disappeared at the delayed posttests.

Instructional Components

There was no demonstrable additional effect of instructional features on immediate measures. Interventions including teacher modeling or individual student practice resulted in similar effect sizes as interventions with only explicit instruction or collaborative activities. However, instructional components affected the delayed effect on comprehension questions, $\Delta\chi^2(4) = 9.39$, $p = .05$. Although teacher modeling did not have a demonstrable effect, we found that when interventions lacked individual activities, students performed much worse on comprehension questions in the delayed posttest ($\Delta g = -1.04$, $SE = 0.39$, $p = .007$). In other words, individual activities worked as a protecting factor against the relapse in scores on delayed posttests. With individual practice during the intervention program, the delayed effect for text structure instruction on comprehension questions would be $g = 0.82$, instead of $g = -0.23$ without individual practice.

TABLE 4
Immediate Effects (Δg) for Content-Related and Instructional Variables

Variable	Outcome measure		
	Comprehension questions	Recall	Summarization
Text structure recognition (intercept)	0.98	1.03	0.22 (ns)
Paragraph-level structure recognition (Δg)		$\Delta 0.57$	$\Delta 0.91$
Structure visualization		$\Delta -0.44$	
Active structure mapping (Δg)	$\Delta 0.39$	$\Delta 0.51$	
Structure-based summarization			
Rule-based summarization (Δg)		$\Delta 0.12$	$\Delta 0.21$
Teacher: No modeling			
Student: No individual practice			

Note. ns = not statistically significant. Empty cells are nonsignificant.

Funnel Plot

Figure 1 shows the funnel plot of the four final models combined. The residuals of the final models with the explanatory variables are plotted against the standard errors. Most points were located in the region between the straight lines. For all outcome measures taken together, only 16 effect sizes (5.7% of the total sample) were identified as outliers. In both the lower bound and upper bound outliers, effect sizes were based on various outcome measures, although comprehension question effect sizes were slightly overrepresented as lower bound outliers. It shows that there was still a statistically significant amount of unexplained heterogeneity in the four data sets: $QE_{\text{Questions}}(82) = 272.78, p < .001$; $QE_{\text{Recall}}(76) = 219.37, p < .001$; $QE_{\text{Summarization}}(45) = 193.31, p < .001$; and $QE_{\text{Knowledge}}(22) = 59.47, p < .001$; but there did not seem to be any systematic heterogeneity.

Discussion

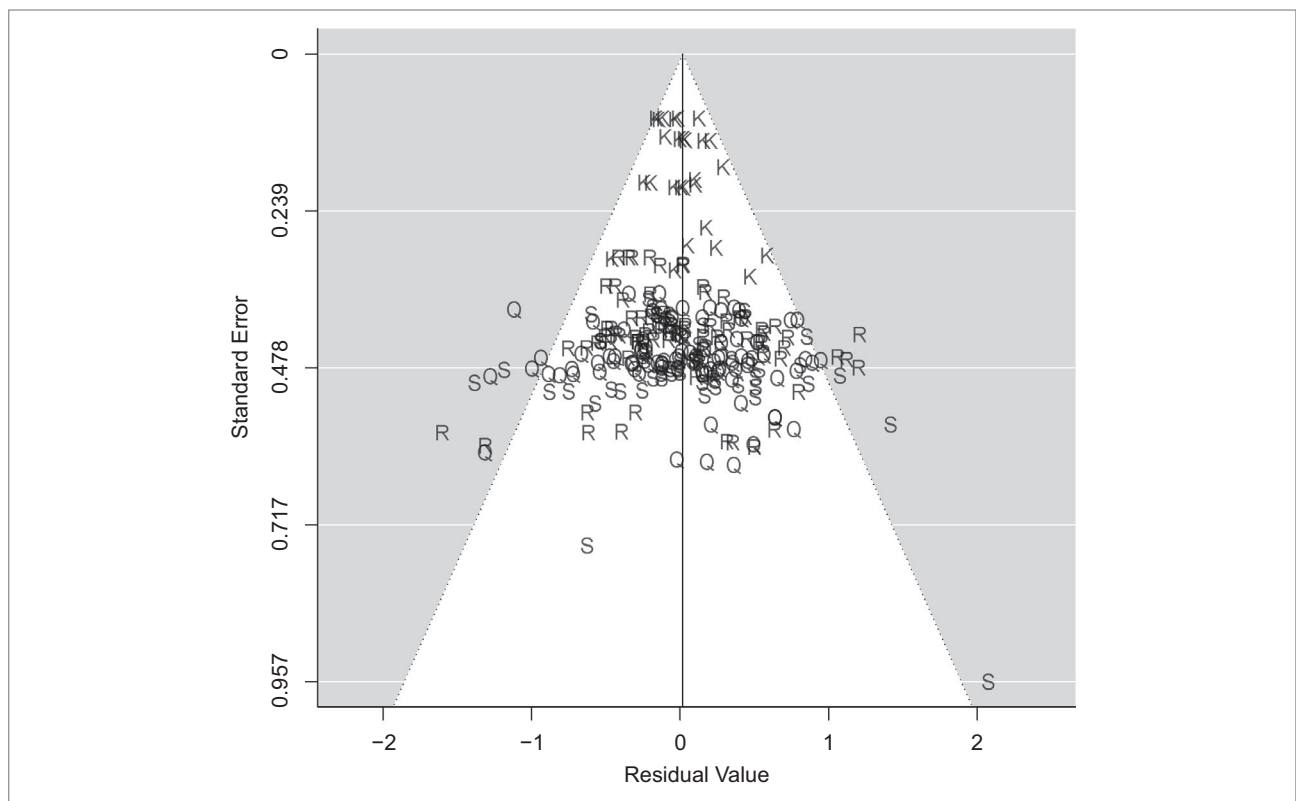
Immediate and Delayed Effects of Text Structure Instruction (Research Question 1)

Our study reveals that text structure instruction can improve reading comprehension skills in grades 4–6.

Narrative and expository interventions seem equally effective in this age group. Contrary to previous meta-analyses (Hebert et al., 2016; Pyle et al., 2017), we could not demonstrate a difference between standardized and nonstandardized tests, but our meta-analysis shows that the type of outcome measure (comprehension questions, recall, summarization, and/or text structure knowledge) has a dramatic impact on effect sizes. When compared with regular reading comprehension instruction, text structure instruction has an overall immediate effect of $g = 0.25$ on comprehension questions, $g = 0.38$ on recall, $g = 0.58$ on summarization, and $g = 0.34$ on text structure knowledge. However, at delayed posttests, the differences between groups that received text structure instruction or regular reading instruction could no longer be demonstrated.

At first glance, the overall effect of text structure instruction on comprehension questions might seem relatively low ($g = 0.25$), but this constitutes an additional effect over and above business-as-usual gains in reading comprehension. Moreover, the effect sizes need to be evaluated in context: They were obtained in authentic educational contexts, not in controlled lab settings, so effect sizes of approximately 0.20 that are often classified as small are actually of policy interest (Durlak, 2009; Hedges & Hedberg, 2007). Furthermore, the effect size on comprehension questions is similar to the effect

FIGURE 1
Funnel Plot of the Final Model



Note. K = text structure knowledge; Q = comprehension questions; R = recall; S = summarization.

of other educational interventions (e.g., $g = 0.24$ in Lipsey et al., 2012). Therefore, text structure instruction should be considered as a way to support the transition to reading-to-learn skills so the persistent fourth-grade slump effect can be reduced (Chall & Jacobs, 1983).

We found that text structure instruction had a similar additional effect when compared with either business-as-usual reading instruction or a heterogeneous subset of alternative interventions (e.g., vocabulary instruction, cognitive reading strategies). This does not imply that, in general, these alternative interventions are not better than business-as-usual instruction. In fact, various meta-analyses have demonstrated the positive effect of vocabulary instruction (Stahl & Fairbanks, 1986) and (meta)cognitive learning strategies (e.g., Donker, de Boer, Kostons, van Ewijk, & van der Werf, 2014) on comprehension. Our result might be due to the fact that the control condition consisted of a very heterogeneous subset of alternative interventions, resulting in a baseline reference group that is comparable to business-as-usual instruction.

The effect size on comprehension questions ($g = 0.25$) was lower than on the other outcome measures; on tests in which students had to apply their knowledge of text structure more directly (e.g., recognizing text structures), the immediate effect size was $g = 0.34$. Although comprehension questions are presented as one outcome measure, they can target various aspects of a text and thereby might still measure different things (Keenan et al., 2008; Nation & Snowling, 1997). Because text structure instruction provides students with the knowledge and tools to process text into more coherent and organized mental schemata (Kintsch, 1988, 2013; Kintsch & van Dijk, 1978), it seems to matter whether comprehension questions involve surface code, textbase, or situation model comprehension skills (Kintsch, 1988, 2013). Unfortunately, most of the studies did not specify whether comprehension questions concerned situation model questions (e.g., main idea questions) or local text issues (e.g., referential coherence) that might be answered without full understanding at the level of the situation model (van den Broek & Kremer, 2000).

Tasks that tap more into situation model comprehension, such as recall and summarization, might be more suitable candidates for evaluating effects of text structure instruction. Indeed, summary and recall measures yielded larger overall effects of text structure instruction ($g = 0.37$ and $.57$, respectively), which is in line with the meta-analytic findings of Pyle and colleagues (2017), who showed that the overall effect of text structure instruction was larger on GO tasks than on comprehension questions. However, even for summarization and recall tasks, it is important to note that they yield qualitatively different results: Summarization tasks typically evoke more main ideas, whereas recall tasks might evoke details as well (Riley & Lee, 1996).

Our meta-analysis shows that outcome measure matters for evaluating the effectiveness of text structure interventions. As each type of outcome measures might rely on a slightly different constellation of comprehension skills and thereby measure different aspects of the reading process (Brown et al., 1983), future text structure research should include multiple outcome measures (Bohaty, Hebert, Nelson, & Brown, 2015), as in the studies by Meyer et al. (2002; Meyer, Wijekumar, & Lin, 2011) and Wijekumar et al. (2012, 2013, 2014). Another important step would be to disentangle how different outcome measures for reading tap into the numerous skills involved in comprehension processes (e.g., vocabulary, strategic knowledge; Graesser et al., 1994; Keenan et al., 2008), so reading interventions can be evaluated more adequately.

An interesting but disappointing finding is the fact that at delayed posttests, the students in the experimental condition no longer outperformed the control group. This resonates with the findings by Hebert and colleagues (2016) that delayed effects of text structure instruction are much smaller and less consistent than immediate effects. In fact, we found that in many studies, the performance of the experimental group decreased between immediate and delayed posttest, whereas the control group performed rather similarly on immediate and delayed posttests or showed a small decrease as well.

A methodological factor that might contribute to this finding is the fact that delayed posttests sometimes required transfer when students were tested on untaught text structures. Due to a limited number of studies with delayed posttests, we could not examine this transfer effect, but Hebert and colleagues (2016) showed much smaller effects in far-transfer cases. Another explanation for the lack of maintenance is the fact that the intervention studies made no effort to promote maintenance; the highlights of text structure instruction were not repeated in the period between the immediate and delayed posttests. Finally, the quality of business-as-usual instruction might be insufficient to help students maintain their newly acquired knowledge about text structures. Teachers in the upper elementary grades often fail to employ evidence-based approaches (Duke et al., 2011; Wijekumar et al., 2019), and struggle themselves with text structure recognition (Beerwinkle et al., 2018; Bogaerds-Hazenberg, Evers-Vermeul, & van den Bergh, 2019; Reutzel et al., 2016) and main idea identification (Kucan et al., 2011). Also, regular textbooks for grades 4 and 5 typically have not included much text structure instruction (Beerwinkle et al., 2018).

Content-Related and Instructional Variables Moderating the Effects (Research Question 2)

The ability to recognize text structures is beneficial for increasing reading comprehension, irrespective of

outcome measure and genre. Our meta-analysis also shows that the effects of moderating content-related and instructional variables are different per outcome measure, which moves the field beyond the question of what elements to include in text structure instruction (e.g., including GOs) by providing some insight into how to include and refine these elements (e.g., how to offer GOs).

Content-Related Variables

When it comes to the effect sizes for summarization and recall, students particularly benefit from text structure instruction that focuses on the paragraph-level structures ($g = 1.60$ on recall and 1.13 on summarization), not just on top-level structures. This corroborates the claim that successful summarization training combines text structure recognition and main idea identification within paragraphs (Miyatsu et al., 2018). The large effect of text structure instruction on summarization supports the hypothesis that text structure provides students with the necessary knowledge and tools to distinguish important from unimportant information (Hogan et al., 2011; Meyer et al., 1980; Taylor, 1985; Winograd, 1984) and that text structure helps students see how these main ideas are organized at a higher level (Miyatsu et al., 2018). Moreover, it can function as a mnemonic aid that improves text recall (Taylor, 1982). Alternatively, it might be easier to establish differences in main ideas at the paragraph level than at the text level, as a paragraph-level focus invites students to produce multiple main ideas instead of a single main idea for the text as a whole. Because all of the interventions included instruction about signaling words, we could not analyze whether a focus on signaling words moderated the effects of text structure instruction.

Structure-based summarization also improves students' text comprehension, especially when students learn specific rules and tricks for paraphrasing the main idea (e.g., rule-based summarization). This yielded net immediate effects of $g = 1.15$ for recall and 0.43 for summarization. This corroborates the idea that explicit knowledge about text structure (e.g., structure-specific questions, signaling words) can provide students with useful tools to identify and formulate main ideas and to reorganize these in a coherent way (Elledge, 2013; Meyer et al., 1980; Miyatsu et al., 2018; E.A. Stevens, 2018; Taylor, 1985) while eliminating redundant information (Brown et al., 1983; McNeil & Donant, 1982). Whereas structure-based summarization techniques were not included as a moderating variable in previous meta-analyses (Hebert et al., 2016; Pyle et al., 2017), our meta-analysis shows that it is an important ingredient of successful text structure interventions. Summarization strategies based on the internal structure of paragraphs seem more helpful in improving students' performance than strategies based on external markers of text structure (i.e., headings and subheadings), possibly because the

former provide students with the necessary skills to distill main ideas from the text, even in the absence of unambiguous external markers of text structure.

Structure visualizations often have been part of larger text structure strategy interventions but also have been used in various studies on GOs and story mapping. Over the years, GO research has generated mixed results, which has often been attributed to the types of GOs used, to the level of instructional support that was provided, or to whether they were used as a prereading or postreading activity (Griffin et al., 1995; Jiang & Grabe, 2007). Although the authors of previous meta-analyses suggested that GOs might increase the effectiveness of text structure interventions (Hebert et al., 2016; Pyle et al., 2017), they did not examine their presence as a moderating variable. Our meta-analysis shows that the inclusion of structure-based visualizations has positive effects on comprehension and recall, as long as students actively fill out these maps and GOs. Simple exposure has no demonstrable effects or even a negative effect on recall. It seems crucial that text structure instruction provides ample opportunities for students to practice filling out structure-based GOs and maps after teacher-led instruction. This underscores the importance of ensuring an instructional approach that is characterized by a gradual release of responsibility from teacher to student (e.g., Fisher & Frey, 2013; Pearson & Gallagher, 1983).

The importance of the active construction of GOs contradicts the conclusion of Stull and Mayer (2007), who found that constructing GOs increased the extraneous cognitive processing load and interfered with learning. In the studies included in our meta-analysis, we found that constructing GOs consisted of students filling out missing information in a teacher-supplied GO and did not require them to draw the whole structure by themselves (i.e., choosing the right boxes, adding arrows). Possibly, this is a less complex task for students, as the text structure is scaffolded in the GO, and students' only concern is to find the right ideas to put in the boxes. This type of task might reduce extraneous load but still fits with the theory that deep learning occurs when students are encouraged to engage in productive learning activities (e.g., Mayer, 2003), so even finishing a partially completed GO provides an opportunity for deep text processing (Jiang & Grabe, 2007). With current technological trends and the development of digital mapping software, it seems relevant to explore digital opportunities for incorporating more learning activities focused on text structure visualizations in classrooms. Although previous meta-analyses showed that teaching multiple text structures has a positive impact on students' performance (Hebert et al., 2016; Pyle et al., 2017), we could not demonstrate this effect. Only for text structure knowledge did we find that students' performance tended to be slightly lower when they encountered more different structures. This seems logical, as students

need to remember more different types of structures on these text structure knowledge tests. Still, teaching multiple structures has at least no demonstrable negative impact on the other comprehension measures. Therefore, we believe that teaching multiple text structures is useful, as students can learn from comparing and contrasting the characteristics of various structures, become more aware of the differences between structures, and possibly even transfer knowledge to untaught text structures (Hebert et al., 2016). Also from a practical perspective, it is important for students to recognize more than one structure, as most (educational) texts are a combination of multiple text structures nested within one another (Jiang & Grabe, 2007). As Pyle et al. (2017) also suggested, we believe that it is worthwhile to engage in further research that addresses the order and complexity of different text structures that are taught.

Instructional Components

Although it matters how a skill is taught to students, most meta-analyses on literacy research have not evaluated the effect of instructional components such as modeling or collaborative practice, possibly because intervention descriptions often extensively focus on content. As Pyle et al. (2017) pointed out, the term *explicit instruction* is often used in describing text structure interventions but is in itself very broad in terms of instructional features present (see also Archer & Hughes, 2011). Despite the limited descriptions of the instructional approach in most research articles, our study still demonstrates that the instructional approach moderates delayed but not immediate effects: Interventions including individual student activities generated slightly higher effects on comprehension questions. Although several instructional models hypothesize practice with peers to be an important step in the gradual release of responsibility from teacher to student (e.g., Fisher & Frey, 2013; Pearson & Gallagher, 1983), we could not demonstrate a moderating effect of collaborative activities. Due to the poor description of instructional features in most of the studies, it is hard to interpret this finding. Students who do not practice alone, only with their peers, may not fully acquire the skill that they have to learn, easily forgetting it and therefore failing on delayed posttests. Alternatively, activities labeled as collaborative in the intervention might not actually have met the criteria for effective cooperation (see Johnson & Johnson, 2017).

Several researchers have expressed the need to pay more attention to the fidelity of implementation in text structure intervention studies (Bohaty et al., 2015). We believe that the quality of instructional components should also be included in such evaluations. More specifically, parallel to the recommendations made by writing researchers (Bouwer & De Smedt, 2018), future reading research articles should systematically provide details on the intervention context and on the design principles of the intervention, at

both a macrolevel (i.e., focus and mode of instruction, sequencing of content) and microlevel (e.g., instructional activities, learning activities, materials). This will increase transparency of intervention results and might also promote implementation of concrete activities in educational contexts (Fidalgo, Harris, & Braaksma, 2018). The gradual release of responsibility model (Fisher & Frey, 2013; Pearson & Gallagher, 1983) provides a useful framework for reading researchers to more systematically describe, test, and evaluate the quality of the instructional components.

Given that we know from various meta-analyses that text structure instruction is effective (the what) for students of various ages (the when), it is important that future studies focus on instructional practice (the how). Now is the time to examine the effectiveness of a greater variety of instructional features in the context of reading instruction so the field can ameliorate the context in which text structure instruction is given (Williams, 2018). A first attempt was already made by Meyer et al. (2002, 2010) and Wijekumar et al. (2012, 2013, 2014), who tried to unravel the effects of providing, for instance, individualized feedback or the effects of tutoring in the context of a web-based intervention. Qualitative research has been undertaken to qualitatively describe teacher instructional practices and pedagogical content knowledge in the context of text structure interventions (Beerwinkle et al., 2018; Bogaerds-Hazenbergh et al., 2019; Wijekumar et al., 2019), which provided more insight into the instructional components that influence intervention success.

In sum, our meta-analysis shows that text structure instruction has a positive effect on students' reading comprehension skills over and above regular reading programs: It improves their performance on comprehension questions, recall, and summarization tasks and has a positive effect on their text structure knowledge. However, at delayed posttests, differences between experimental and control groups can no longer be demonstrated. Hence, it seems a promising avenue to incorporate text structure instruction into primary school curricula so students' comprehension skills can be strengthened and positive effects maintained over time.

NOTE

This work was supported by a Research Talent grant (project 406.16.052) by the Netherlands Organization for Scientific Research (NWO).

REFERENCES

- References marked with an asterisk indicate studies included in the meta-analysis.
- Alvermann, D.E., & Boothby, P.R. (1984, April). *Knowledge of text structure and its influence on a transfer task*. Paper presented at the 68th annual meeting of the American Educational Research Association, New Orleans, LA.

- *Alvermann, D.E., & Boothby, P.R. (1986). Children's transfer of graphic organizer instruction. *Reading Psychology*, 7(2), 87–100. <https://doi.org/10.1080/0270271860070203>
- Anderson, R.C., & Pearson, P.D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P.D. Pearson, R. Barr, M.L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 225–291). New York, NY: Longman.
- Archer, A., & Hughes, C.A. (2011). *Explicit instruction: Efficient and effective teaching*. New York, NY: Guilford.
- Armbruster, B.B., & Anderson, T.H. (1982). *Idea-mapping: The technique and its use in the classroom or simulating the “ups” and “downs” of reading comprehension* (Reading Education Report No. 36). Champaign: Center for the Study of Reading, University of Illinois at Urbana-Champaign; Cambridge, MA: Bolt Beranek and Newman.
- *Armbruster, B.B., Anderson, T.H., & Meyer, J.L. (1991). Improving content-area reading using instructional graphics. *Reading Research Quarterly*, 26(4), 393–416. <https://doi.org/10.2307/747895>
- *Armbruster, B.B., Anderson, T.H., & Ostertag, J. (1987). Does text structure/summarization instruction facilitate learning from expository text? *Reading Research Quarterly*, 22(3), 331–346. <https://doi.org/10.2307/747972>
- Barron, R.F. (1980, December). *A systematic research procedure, organizers, and overviews: An historical perspective*. Paper presented at the annual meeting of the National Reading Conference, San Diego, CA.
- Bartlett, B.J. (1978). *Top-level structure as an organizational strategy for recall of classroom text*. Unpublished doctoral dissertation, Arizona State University, Tempe.
- Baumann, J.F. (1983). Children's ability to comprehend main ideas in content textbooks. *Literacy Research and Instruction*, 22(4), 322–331. <https://doi.org/10.1080/19388078309557722>
- *Bean, T.W., & Steenwyk, F.L. (1984). The effect of three forms of summarization instruction on sixth graders' summary writing and comprehension. *Journal of Reading Behavior*, 16(4), 297–306. <https://doi.org/10.1080/10862968409547523>
- Beerwinkle, A.L., Wijekumar, K., Walpole, S., & Aguis, R. (2018). An analysis of the ecological components within a text structure intervention. *Reading and Writing*, 31(9), 2041–2064. <https://doi.org/10.1007/s11145-018-9870-5>
- *Berkowitz, S.J. (1986). Effects of instruction in text organization on sixth-grade students' memory for expository reading. *Reading Research Quarterly*, 21(2), 161–178. <https://doi.org/10.2307/747843>
- Bogaerds-Hazenberg, S.T.M., Evers-Vermeul, J., & van den Bergh, H. (2019). Teachers and researchers as co-designers? A design-based research on reading comprehension instruction in primary education. *Educational Design Research*, 3(1), article 18. <https://doi.org/10.15460/eder.3.1.1399>
- *Bohaty, J.J. (2015). *The effects of expository text structure instruction on the reading outcomes of 4th and 5th graders experiencing reading difficulties* (Doctoral dissertation). Retrieved from <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1230&context=cehdsiss>
- Bohaty, J.J., Hebert, M.A., Nelson, J.R., & Brown, J.A. (2015). Methodological status and trends in expository text structure instruction efficacy research. *Reading Horizons*, 54(2), 35–65.
- *Boothby, P.R., & Alvermann, D.E. (1984). A classroom training study: The effects of graphic organizer instruction on fourth graders' comprehension. *Reading World*, 23(4), 325–339. <https://doi.org/10.1080/19388078409557783>
- Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.
- Bouwer, R., & De Smedt, F. (2018). Guidelines for reporting instructional writing interventions in journal articles [Editorial]. *Journal of Writing Research*, 10(2), 115–137. <https://doi.org/10.17239/jowr-2018.10.02.01>
- *Bowman, M., & Gambrell, L. (1981, April). *The effects of story structure questioning upon reading comprehension*. Paper presented at the 54th annual meeting of the American Educational Research Association, Los Angeles, CA.
- *Braxton, D.M. (2009). *The effects of two summarization strategies using expository text on the reading comprehension and summary writing of fourth- and fifth-grade students in an urban, Title 1 school* (Doctoral dissertation). Retrieved from <https://drum.lib.umd.edu/handle/1903/9918>
- *Broer, N.A., Aarnoutse, C.A.J., Kieviet, F.K., & van Leeuwe, J.F.J. (2002). The effect of instructing the structural aspect of texts. *Educational Studies*, 28(3), 213–238. <https://doi.org/10.1080/0305569022000003681>
- Brown, A.L., Day, J.D., & Jones, R.S. (1983). Macrorules for summarizing texts: The development of plans for summarizing texts. *Child Development*, 54(4), 968–979. <https://doi.org/10.2307/1129901>
- Bus, A.G., & van IJzendoorn, M.H. (1999). Phonological awareness in early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, 91(3), 403–414. <https://doi.org/10.1037/0022-0663.91.3.403>
- *Carriedo, N., & Alonso-Tapia, J. (1996). Main idea comprehension: Training teachers and effects on students. *Journal of Research in Reading*, 19(2), 128–153. <https://doi.org/10.1111/j.1467-9817.1996.tb00093.x>
- Chall, J.S., & Jacobs, V.A. (1983). Writing and reading in the elementary grades: Developmental trends among low SES children. *Language Arts*, 60(5), 617–626.
- Chall, J.S., & Jacobs, V.A. (2003). Poor children's fourth-grade slump. *American Educator*, 27(1), 14–44.
- Ciullo, S., & Reutebuch, C. (2013). Computer-based graphic organizers for students with LD: A systematic review of literature. *Learning Disabilities Research & Practice*, 28(4), 196–210. <https://doi.org/10.1111/ldrp.12017>
- Cohen, E.G., & Benton, J. (1988). Making groupwork work. *American Educator*, 12(3), 10–17.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Sage.
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.
- *Darch, C.B., Carnine, D.W., & Kameenui, E.J. (1986). The role of graphic organizers and social structure in content area instruction. *Journal of Reading Behavior*, 18(4), 275–295. <https://doi.org/10.1080/10862968609547576>
- *de Jou, G.I., & Sperb, T.M. (2009). Lectura comprensiva: Un estudio de intervención/Comprehensive reading: An intervention study. *Revista Interamericana Journal de Psicología*, 43(1), 12–21.
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students: A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231–264. <https://doi.org/10.1007/s11409-008-9029-x>
- Donker, A.S., de Boer, H., Kostons, D., van Ewijk, C.C.D., & van der Werf, M.P.C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, 11, 1–26. <https://doi.org/10.1016/j.edurev.2013.11.002>
- Duffy, G.G. (2002). The case for direct explanation of strategies. In C.C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 28–41). New York, NY: Guilford.
- Duke, N.K. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly*, 35(2), 202–224. <https://doi.org/10.1598/RRQ.35.2.1>
- Duke, N.K., Pearson, P.D., Strachan, S.L., & Billman, A.K. (2011). Essential elements of fostering and teaching reading comprehension.

- In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about reading instruction* (4th ed., pp. 51–93). Newark, DE: International Reading Association.
- Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly*, 14(4), 481–533. <https://doi.org/10.1598/RRQ.14.4.2>
- Durlak, J.A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology*, 34(9), 917–928. <https://doi.org/10.1093/jpepsy/jsp004>
- *Elledge, D.H. (2013). *Improving reading comprehension through explicit summarization instruction* (Doctoral dissertation). Retrieved from https://etd.ohiolink.edu/etd.send_file?accession=ucin1367928451&disposition=inline
- *Ermis, S. (2008). Using graphic organizers to facilitate elementary students' comprehension of informational text. In M.M. Foote, F. Falk-Ross, S. Szabo, & M.B. Sampson (Eds.), *Navigating the literacy waters: Research, praxis, and advocacy* (pp. 87–102). Logan, UT: College Reading Association.
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen. (2009). *Referentiekader taal en rekenen: De referentieniveaus* [Framework of reference for language and arithmetic: The reference levels]. Enschede, The Netherlands: Doorlopende Leerlijnen Taal en Rekenen.
- Fidalgo, R., Harris, K.R., & Braaksma, M. (Eds.). (2018). *Design principles for teaching effective writing: Theoretical and empirical grounded principles*. Leiden, The Netherlands: Brill.
- Fisher, D., & Frey, N. (2013). *Better learning through structured teaching: A framework for the gradual release of responsibility* (2nd ed.). Alexandria, VA: ASCD.
- *Fitzgerald, J., & Spiegel, D.L. (1983). Enhancing children's reading comprehension through instruction in narrative structure. *Journal of Literacy Research*, 15(2), 1–17. <https://doi.org/10.1080/10862968309547480>
- Frey, N., Fisher, D., & Hernandez, T. (2003). "What's the gist?": Summary writing for struggling adolescent writers. *Voices From the Middle*, 11(2), 43–49.
- Fritz, C.O., Morris, P.E., & Richler, J.J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1), 2–18. <https://doi.org/10.1037/a0024338>
- Fuchs, D., Fuchs, L.S., Mathes, P.G., & Simmons, D.C. (1997). Peer-Assisted Learning Strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34(1), 174–206. <https://doi.org/10.3102/00028312034001174>
- Gardill, M.C., & Jitendra, A.K. (1999). Advanced story map instruction: Effects on the reading comprehension of students with learning disabilities. *The Journal of Special Education*, 33(1), 2–17. <https://doi.org/10.1177/002246699903300101>
- Garibaldi, A.M. (1979). Affective contributions of cooperative and group goal structures. *Journal of Educational Psychology*, 71(6), 788–794. <https://doi.org/10.1037/0022-0663.71.6.788>
- Garner, R. (1987). Strategies for reading and studying expository text. *Educational Psychologist*, 22(3/4), 299–312. <https://doi.org/10.1080/00461520.1987.9653054>
- *Gentry, L.J. (2006). *Comparison of the effects of training in expository text structure through annotation text marking and training in vocabulary development on reading comprehension of students going into fourth grade* (Doctoral dissertation). Retrieved from <https://digital.lib.usf.edu/SFS0025739/00001>
- Gersten, R., Fuchs, L.S., Williams, J.P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, 71(2), 279–320. <https://doi.org/10.3102/00346543071002279>
- Goldman, S.R. (1997). Learning from text: Reflections on the past and suggestions for the future. *Discourse Processes*, 23(3), 357–398. <https://doi.org/10.1080/01638539709544997>
- *Gordon, C.J., & Pearson, P.D. (1983). *The effects of instruction in meta-comprehension and inferencing on children's comprehension abilities* (Technical Report No. 277). Washington, DC: National Institute of Education.
- Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6–10. <https://doi.org/10.1177/074193258600700104>
- Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371–395. <https://doi.org/10.1037/0033-295X.101.3.371>
- *Griffin, C.C., Malone, L.D., & Kameenui, E.J. (1995). Effects of graphic organizer instruction on fifth-grade students. *The Journal of Educational Research*, 89(2), 98–107. <https://doi.org/10.1080/00220671.1995.9941200>
- Hansen, C.L. (1978). Story retelling used with average and learning disabled readers as a measure of reading comprehension. *Learning Disability Quarterly*, 1(3), 62–69. <https://doi.org/10.2307/1510938>
- Hare, V.C., & Borhardt, K.M. (1984). Direct instruction of summarization skills. *Reading Research Quarterly*, 20(1), 62–78.
- Haria, P.D., & Midgett, E. (2014). A genre-specific reading comprehension strategy to enhance struggling fifth-grade readers' ability to critically analyze argumentative text. *Reading & Writing Quarterly*, 30(4), 297–327. <https://doi.org/10.1080/10573569.2013.818908>
- Hebert, M., Bohaty, J.J., Nelson, J.R., & Brown, J. (2016). The effects of text structure instruction on expository reading comprehension: A meta-analysis. *Journal of Educational Psychology*, 108(5), 609–629. <https://doi.org/10.1037/edu0000082>
- Hedges, L.V., & Hedberg, E.C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, 29(1), 60–87. <https://doi.org/10.3102/0162373707299706>
- Hiebert, E.H., & Mesmer, H.A.E. (2013). Upping the ante of text complexity in the Common Core State Standards: Examining its potential impact on young readers. *Educational Researcher*, 42(1), 44–51. <https://doi.org/10.3102/0013189X12459802>
- *Hoffman, K.F. (2010). *The impact of graphic organizer and metacognitive monitoring instruction on expository science text comprehension in fifth grade students* (Doctoral dissertation). Retrieved from <https://repository.lib.ncsu.edu/handle/1840.16/6198>
- Hogan, T.P., Bridges, M.S., Justice, L.M., & Cain, K. (2011). Increasing higher level language skills to improve reading comprehension. *Focus on Exceptional Children*, 44(3), 1–18, 20. <https://doi.org/10.17161/foec.v44i3.6688>
- Idol, L., & Croll, V.J. (1987). Story-mapping training as a means of improving reading comprehension. *Learning Disability Quarterly*, 10(3), 214–229. <https://doi.org/10.2307/1510494>
- Ishikawa, K. (1971). *Guide to quality control*. Tokyo, Japan: Asian Productivity Organization.
- Jiang, X., & Grabe, W. (2007). Graphic organizers in reading instruction: Research findings and issues. *Reading in a Foreign Language*, 19(1), 34–55.
- Johnson, D.W., & Johnson, F.P. (2017). *Joining together: Group theory and group skills* (12th ed.). Boston, MA: Pearson.
- Jones, B.F., Pierce, J., & Hunter, B. (1988). Teaching students to construct graphic representations. *Educational Leadership*, 46(4), 20–25.
- Keenan, J.M., Betjemann, R.S., & Olson, R.K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, 12(3), 281–300. <https://doi.org/10.1080/10888430802132279>
- Kelly, C.H. (2019). A backwards glance: How 20th century reading research and practices have shaped current instructional approaches. *Making Literacy Connections*, 33, 4–12.
- Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, 35(7), 1567–1577. <https://doi.org/10.3758/BF03193491>

- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163–182. <https://doi.org/10.1037/0033-295X.95.2.163>
- Kintsch, W. (2013). Revisiting the construction-integration model of text comprehension and its implications for instruction. In D.E. Alvermann, N.J. Unrau, & R.B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 807–839). Newark, DE: International Reading Association.
- Kintsch, W., & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394. <https://doi.org/10.1037/0033-295X.85.5.363>
- Kostons, D., Donker, A.S., & Opdenakker, M.C. (2014). *Zelfgestuurd leren in de onderwijspraktijk: Een kennisbasis voor effectieve strategie-instructie* [Self-regulated learning: A knowledge base for effective strategy instruction]. Groningen, The Netherlands: GION.
- Kucan, L., Hapgood, S., & Palincsar, A.S. (2011). Teachers' specialized knowledge for supporting student comprehension in text-based discussions. *The Elementary School Journal*, 112(1), 61–82. <https://doi.org/10.1086/660689>
- Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior*, 25(2), 284–289. <https://doi.org/10.1016/j.chb.2008.12.010>
- Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., ... Busick, M.D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms* (NCSE 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.
- Mandler, J.M., & Johnson, N.S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9(1), 111–151. [https://doi.org/10.1016/0010-0285\(77\)90006-8](https://doi.org/10.1016/0010-0285(77)90006-8)
- Manoli, P., & Papadopoulou, M. (2012). Graphic organizers as a reading strategy: Research findings and issues. *Creative Education*, 3(3), 348–356. <https://doi.org/10.4236/ce.2012.33055>
- Mayer, R.E. (2003). *Learning and instruction*. Upper Saddle River, NJ: Merrill.
- *McDermott, M.B. (1991). *Teaching top level structure as an aid for reading comprehension of expository prose for fourth-grade pupils*. Unpublished doctoral dissertation, Hofstra University, Hempstead, NY.
- McKeown, M.G., Beck, I.L., & Blake, R.G.K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly*, 44(3), 218–253. <https://doi.org/10.1598/RRQ.44.3.1>
- *McLaughlin, E.M. (1990). *Effects of graphic organizers and levels of text difficulty on less-proficient fifth-grade readers' comprehension of expository text*. Unpublished doctoral dissertation, University of Maryland, College Park.
- McNeil, J., & Donant, L. (1982). Summarization strategy for improving reading comprehension. In J.A. Niles & L.A. Harris (Eds.), *New inquiries in reading research and instruction: Thirty-first yearbook of the National Reading Conference* (pp. 215–219). Rochester, NY: National Reading Conference.
- Meyer, B.J.F. (1975). Identification of the structure of prose and its implications for the study of reading and memory. *Journal of Reading Behavior*, 7(1), 7–47. <https://doi.org/10.1080/10862967509547120>
- Meyer, B.J.F., Brandt, D.M., & Bluth, G.J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, 16(1), 72–103. <https://doi.org/10.2307/747349>
- Meyer, B.J.F., & Freedle, R.O. (1984). Effects of discourse type on recall. *American Educational Research Journal*, 21(1), 121–143. <https://doi.org/10.3102/00028312021001121>
- *Meyer, B.J.F., Middlemiss, W., Theodorou, E., Brezinski, K.L., McDougall, J., & Bartlett, B.J. (2002). Effects of structure strategy instruction delivered to fifth-grade children using the internet with and without the aid of older adult tutors. *Journal of Educational Psychology*, 94(3), 486–519. <https://doi.org/10.1037/0022-0663.94.3.486>
- Meyer, B.J.F., & Ray, M.N. (2011). Structure strategy interventions: Increasing reading comprehension of expository text. *International Electronic Journal of Elementary Education*, 4(1), 127–152.
- Meyer, B.J.F., & Rice, G.E. (1984). The structure of text. In P.D. Pearson, R. Barr, M.L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 319–351). New York, NY: Longman.
- *Meyer, B.J.F., Wijekumar, K.K., & Lin, Y.-C. (2011). Individualizing a web-based structure strategy intervention for fifth graders' comprehension of nonfiction. *Journal of Educational Psychology*, 103(1), 140–168. <https://doi.org/10.1037/a0021606>
- Meyer, B.J.F., Wijekumar, K., Middlemiss, W., Higley, K., Lei, P.-W., Meier, C., & Spielvogel, J. (2010). Web-based tutoring of the structure strategy with or without elaborated feedback or choice for fifth- and seventh-grade readers. *Reading Research Quarterly*, 45(1), 62–92. <https://doi.org/10.1598/RRQ.45.1.4>
- Miyatsu, T., Nguyen, K., & McDaniel, M.A. (2018). Five popular study strategies: Their pitfalls and optimal implementations. *Perspectives on Psychological Science*, 13(3), 390–407. <https://doi.org/10.1177/1745691617710510>
- *Moore, S.R. (1996, April). *Collaboration and the reading-writing relationship: Implications for building schemata for expository text*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.
- Moorf, D.W., & Readence, J.E. (1984). A quantitative and qualitative review of graphic organizer research. *The Journal of Educational Research*, 78(1), 11–17. <https://doi.org/10.1080/00220671.1984.10885564>
- Moss, B., & Newton, E. (2002). An examination of the informational text genre in basal readers. *Reading Psychology*, 23(1), 1–13. <https://doi.org/10.1080/027027102317345376>
- Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, 67(3), 359–370. <https://doi.org/10.1111/j.2044-8279.1997.tb01250.x>
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.
- Nolte, R.Y., & Singer, H. (1985). Active comprehension: Teaching a process of reading comprehension and its effects on reading achievement. *The Reading Teacher*, 39(1), 24–31.
- Oakhill, J., Cain, K., & Elbro, C. (2015). *Understanding and teaching reading comprehension: A handbook*. London, UK: Routledge.
- *Ocasio, T.L. (2006). *A comparison of two instructional programs to develop strategies to improve reading comprehension* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3209082)
- Ogle, D., & Blachowicz, C.L. (2002). Beyond literature circles: Helping students comprehend informational texts. In C.C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 259–274). New York, NY: Guilford.
- Pearson, P.D., & Gallagher, M.C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, 8(3), 317–344. [https://doi.org/10.1016/0361-476X\(83\)90019-X](https://doi.org/10.1016/0361-476X(83)90019-X)
- *Ponce, H.R., López, M.J., & Mayer, R.E. (2012). Instructional effectiveness of a computer-supported program for teaching reading comprehension strategies. *Computers & Education*, 59(4), 1170–1183. <https://doi.org/10.1016/j.compedu.2012.05.013>
- Pyle, N., Vasquez, A.C., Lignugaris/Kraft, B., Gillam, S.L., Reutzell, D.R., Olszewski, A., ... Pyle, D. (2017). Effects of expository text structure interventions on comprehension: A meta-analysis. *Reading Research Quarterly*, 52(4), 469–501. <https://doi.org/10.1002/rrq.179>

- *Raphael, T.E., Englert, C.S., & Kirschner, B.W. (1986). *The impact of text structure instruction and social context on students' comprehension and production of expository text* (Research Series No. 177). East Lansing: Institute for Research on Teaching, Michigan State University.
- *Raphael, T.E., & Kirschner, B.M. (1985). *The effects of instruction in compare/contrast text structure on sixth-grade students' reading comprehension and writing products* (Research Series No. 161). East Lansing: Institute for Research on Teaching, Michigan State University.
- Rapp, D.N., van den Broek, P., McMaster, K.L., 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), 289–312. <https://doi.org/10.1080/10888430701530417>
- Reutzel, D.R., Jones, C.D., Clark, S.K., & Kumar, T. (2016). The Informational Text Structure Survey (ITS²): An exploration of primary grade teachers' sensitivity to text structure in young children's informational texts. *The Journal of Educational Research, 109*(1), 81–98. <https://doi.org/10.1080/00220671.2014.918927>
- Riley, G.L., & Lee, J.F. (1996). A comparison of recall and summary protocols as measures of second language reading comprehension. *Language Testing, 13*(2), 173–189. <https://doi.org/10.1177/026553229601300203>
- Robinson, D.H., & Molina, E. (2002). The relative involvement of visual and auditory working memory when studying adjunct displays. *Contemporary Educational Psychology, 27*(1), 118–131. <https://doi.org/10.1006/ceps.2001.1081>
- Robinson, D.H., & Skinner, C.H. (1996). Why graphic organizers facilitate search processes: Fewer words or computationally efficient indexing? *Contemporary Educational Psychology, 21*(2), 166–180. <https://doi.org/10.1006/ceps.1996.0014>
- Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*(4), 479–530. <https://doi.org/10.3102/00346543064004479>
- Sanders, T.J.M., & Noordman, L.G.M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes, 29*(1), 37–60. https://doi.org/10.1207/S15326950dp2901_3
- Sanders, T., & Spooren, W. (2007). Discourse and text structure. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 916–941). Oxford, UK: Oxford University Press.
- 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* (Vol. 1, pp. 97–110). New York, NY: Guilford.
- *Scott, D.B. (2011). *Explicit instruction on rhetorical patterns and student-constructed graphic organizers: The impact on sixth-grade students' comprehension of social studies text* (Doctoral dissertation). Retrieved from <https://drum.lib.umd.edu/handle/1903/11736>
- *Short, E.J., & Ryan, E.B. (1984). Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. *Journal of Educational Psychology, 76*(2), 225–235. <https://doi.org/10.1037/0022-0663.76.2.225>
- Singer, H., & Donlan, D. (1980). *Reading and learning from text*. Boston, MA: Little, Brown.
- Slavin, R.E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research, 57*(3), 293–336. <https://doi.org/10.3102/00346543057003293>
- Smith, S.J., & Okolo, C. (2010). Response to Intervention and evidence-based practices: Where does technology fit? *Learning Disability Quarterly, 33*(4), 257–272. <https://doi.org/10.1177/073194871003300404>
- Spiegel, G.F., & Barufaldi, J.P. (1994). The effects of a combination of text structure awareness and graphic postorganizers on recall and retention of science knowledge. *Journal of Research in Science Teaching, 31*(9), 913–932. <https://doi.org/10.1002/tea.3660310907>
- *Spires, H.A., Gallini, J., & Riggsbee, J. (1992). Effects of schema-based and text structure-based cues on expository prose comprehension in fourth graders. *Journal of Experimental Education, 60*(4), 307–320. <https://doi.org/10.1080/00220973.1992.9943868>
- Stahl, S.A., & Fairbanks, M.M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*(1), 72–110. <https://doi.org/10.3102/00346543056001072>
- *Stevens, E.A. (2018). *The effects of a text structure and paraphrasing intervention on the main idea generation and reading comprehension of struggling readers in grades 4 and 5* (Doctoral dissertation). Retrieved from <https://repositories.lib.utexas.edu/handle/2152/69020>
- Stevens, R.J., Slavin, R.E., & Farnish, A.M. (1991). The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology, 83*(1), 8–16. <https://doi.org/10.1037/0022-0663.83.1.8>
- Stine-Morrow, E.A., Gagne, D.D., Morrow, D.G., & DeWall, B.H. (2004). Age differences in rereading. *Memory & Cognition, 32*(5), 696–710. <https://doi.org/10.3758/BF03195860>
- Stull, A.T., & Mayer, R.E. (2007). Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology, 99*(4), 808–820. <https://doi.org/10.1037/0022-0663.99.4.808>
- *Tackett, S.A., Patberg, J.P., & Dewitz, P. (1984). The effects of story structure instruction on the recall of poor sixth grade readers. *Reading Psychology, 5*(3), 193–201. <https://doi.org/10.1080/0270271840050303>
- *Taylor, B.M. (1982). Text structure and children's comprehension and memory for expository material. *Journal of Educational Psychology, 74*(3), 323–340. <https://doi.org/10.1037/0022-0663.74.3.323>
- *Taylor, B.M. (1985). Improving middle-grade students' reading and writing of expository text. *The Journal of Educational Research, 79*(2), 119–125. <https://doi.org/10.1080/00220671.1985.10885661>
- Taylor, B.M., & Beach, R.W. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly, 19*(2), 134–146.
- *Troyer, S.J. (1994, April). *The effects of three instructional conditions in text structure on upper elementary students' reading comprehension and writing performance*. Paper presented at the 75th annual meeting of the American Educational Research Association, New Orleans, LA.
- Turcotte, C., Giguère, M.-H., & Godbout, M.-J. (2015). Une approche d'enseignement des stratégies de compréhension de lecture de textes courants auprès de jeunes lecteurs à risque d'échouer [An instructional approach to teach reading strategies to young readers at risk for reading failure]. *Language & Literacy, 17*(1), 106–125. <https://doi.org/10.20360/G2SW2B>
- *Ulper, H., & Akkok, E.A. (2010). The effect of using expository text structures as a strategy on summarization skills. In L.E. Kattington (Ed.), *Handbook of curriculum development* (pp. 303–328). New York, NY: Nova Science.
- van den Broek, P., & Kremer, K.E. (2000). The mind in action: What it means to comprehend during reading. In B.M. Taylor, M.F. Graves, & P. van den Broek (Eds.), *Reading for meaning: Fostering comprehension in the middle grades* (pp. 1–31). New York, NY: Teachers College Press; Newark, DE: International Reading Association.
- van der Schoot, M., Horsley, T.M., & van Lieshout, E.C.D.M. (2010). The effects of instruction on situation model construction: An eye fixation study on text comprehension in primary school children. *Educational Psychology, 30*(7), 817–835. <https://doi.org/10.1080/01443410.2010.510600>
- van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic.
- van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22*(2), 155–174. <https://doi.org/10.1007/s10648-010-9134-7>
- van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative

- and expository texts. *Discourse Processes*, 52(1), 47–76. <https://doi.org/10.1080/0163853X.2014.905237>
- *Van Steenbrugge, H. (2006). *Begrijpend lezen en schematiseren: impact van schema's maken op metacognitie* [Reading comprehension and schematizing: Impact of schematizing on metacognition] (Master's thesis). Retrieved from <https://lib.ugent.be/en/catalog/rug01:001043530>
- Vaughn, S., Klingner, J.K., & Bryant, D.P. (2001). Collaborative strategic reading as a means to enhance peer-mediated instruction for reading comprehension and content-area learning. *Remedial and Special Education*, 22(2), 66–74. <https://doi.org/10.1177/074193250102200201>
- *Vidal-Abarca, E. (1990). Un programa para la enseñanza de la comprensión de ideas principales de textos expositivos/A program for teaching comprehension of main ideas in expository texts. *Journal for the Study of Education and Development*, 13(49), 53–71. <https://doi.org/10.1080/02103702.1990.10822257>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- *Walker, M.L. (1995). Help for the “fourth-grade slump”—SRQ2R plus instruction in text structure or main idea. *Reading Horizons*, 36(1), 38–58.
- Weaver, P.A., & Dickinson, D.K. (1982). Scratching below the surface structure: Exploring the usefulness of story grammar. *Discourse Processes*, 5(3/4), 225–243. <https://doi.org/10.1080/01638538209544544>
- Westby, C., Culatta, B., Lawrence, B., & Hall-Kenyon, K. (2010). Summarizing expository texts. *Topics in Language Disorders*, 30(4), 275–287. <https://doi.org/10.1097/TLD.0b013e3181ff5a88>
- Wijekumar, K., Beerwinkle, A.L., Harris, K.R., & Graham, S. (2019). Etiology of teacher knowledge and instructional skills for literacy at the upper elementary grades. *Annals of Dyslexia*, 69(1), 5–20. <https://doi.org/10.1007/s11881-018-00170-6>
- *Wijekumar, K.K., Meyer, B.J.F., & Lei, P. (2012). Large-scale randomized controlled trial with 4th graders using intelligent tutoring of the structure strategy to improve nonfiction reading comprehension. *Educational Technology Research and Development*, 60(6), 987–1013. <https://doi.org/10.1007/s11423-012-9263-4>
- *Wijekumar, K., Meyer, B.J.F., & Lei, P. (2013). High-fidelity implementation of web-based intelligent tutoring system improves fourth and fifth graders content area reading comprehension. *Computers & Education*, 68(1), 366–379. <https://doi.org/10.1016/j.compedu.2013.05.021>
- *Wijekumar, K., Meyer, B.J.F., Lei, P.-W., Lin, Y.-C., Johnson, L.A., Spielvogel, J.A., ... Cook, M. (2014). Multisite randomized controlled trial examining intelligent tutoring of structure strategy for fifth-grade readers. *Journal of Research on Educational Effectiveness*, 7(4), 331–357. <https://doi.org/10.1080/19345747.2013.853333>
- Williams, J.P. (2018). Text structure instruction: The research is moving forward. *Reading and Writing*, 31(9), 1923–1935. <https://doi.org/10.1007/s11145-018-9909-7>
- Winograd, P.N. (1984). Strategic difficulties in summarizing texts. *Reading Research Quarterly*, 19(4), 404–425. <https://doi.org/10.2307/747913>
- Wolman, C. (1991). Sensitivity to causal cohesion in stories by children with mild mental retardation, children with learning disabilities, and children without disabilities. *The Journal of Special Education*, 25(2), 135–154. <https://doi.org/10.1177/002246699102500202>
- Zwaan, R.A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 920–933. <https://doi.org/10.1037/0278-7393.20.4.920>
- Zwaan, R.A., & Radvansky, G.A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–185. <https://doi.org/10.1037/0033-2909.123.2.162>

Submitted November 9, 2018

Final revision received February 7, 2020

Accepted February 11, 2020

SUZANNE T.M. BOGAERDS-HAZENBERG

(corresponding author) is a doctoral student at the Utrecht Institute of Linguistics OTS at Utrecht University, The Netherlands; email s.t.m.bogaerds-hazenberg@uu.nl. Her research focuses on text structure instruction in the upper elementary grades.

JACQUELINE EVERS-VERMEUL is an assistant professor in the Department of Languages, Literature and Communication at Utrecht University, The Netherlands; email j.evers@uu.nl. She studies the representation, production, comprehension, and teaching of text-structural aspects of discourse.

HUUB VAN DEN BERGH is a full professor in the Department of Languages, Literature and Communication at Utrecht University, The Netherlands; email h.vandenbergh@uu.nl. His main research focuses on the didactics and assessment of language teaching in various contexts.

APPENDIX

Model-Fitting Procedures

Here, we provide an overview of our stepwise model fitting and the parameters of the final model for each of the outcome measures. Because the analysis of moderating content-related and instructional variables (models 6 and 7 for comprehension questions and recall, and models 5 and 6 for summarization) might lead to an excessive number of cells, in a second step, we compared this full-factorial model with a more economic model including only the dummy variables that seemed to matter and had sufficient independent observations in our data sets. Specifically, we first

analyzed whether the immediate and delayed effects of text structure instruction were affected by structure visualizations (e.g., maps, GOs) and structure-based summary techniques (e.g., main idea identification, outlining). This did not improve the model fit: comprehension questions, $\Delta\chi^2(2) = 0.73, p = .69$; recall, $\Delta\chi^2(2) = 0.082, p = .96$; and summarization, $\Delta\chi^2(2) = 5.78, p = .06$.

Because the text structure interventions were very heterogeneous, we added three more specific dummy variables: a focus on paragraph structure (i.e., topic

sentences, supporting details), active use of GOs, and teaching rule-based summarization. This improved model fit for all outcome measures: comprehension questions, $\Delta\chi^2(2) = 6.57, p = .04$; recall, $\Delta\chi^2(5) = 15.49, p = .002$; and summarization, $\Delta\chi^2(5) = 15.49, p = .002$. For recall, we could not simplify the full factorial model because almost all parameter estimates were statistically significant. For the other outcome measures, we concluded that the simplified model had no reduced model fit when compared with the full factorial model: comprehension

questions, $\Delta\chi^2(3) = 1.46, p = .69$; and summarization, $\Delta\chi^2(2) = 0.79, p = .67$. Subsequently, we added two instructional features to the model: teacher modeling (in addition to or instead of explicit instruction only) and individual student activities (in addition to or instead of collaborative activities only), both also in interaction with measurement timepoint. Adding these instructional components only improved the model fit for comprehension questions, $\Delta\chi^2(4) = 9.39, p = .05$, not for the other outcome measures.

Comprehension Questions

Overview of Stepwise Model Fitting for Comprehension Questions

Model (M)	-2 log likelihood ratio	Comparison	Δx^2	Δdf	<i>p</i>
M0	145.10				
M1: M0 + Δ Publication	143.78	M0 vs. M1	1.32	1	.24
M2: M0 + Δ Competing_Intervention	143.92	M0 vs. M2	1.79	1	.28
M3: M0 + Δ Timepoint	138.22	M0 vs. M3	1.79	1	.009
M4a: M3 + Δ Genre	138.20	M3 vs. M4a	0.12	1	.73
M4b: M3 + Δ Genre + (Δ Genre \times Timepoint)	137.20	M3 vs. M4b	1.01	2	.60
M5a: M3 + Δ SN	137.67	M3 vs. M5a	0.54	1	.46
M5b: M3 + Δ SN + (Δ SN \times Timepoint)	134.77	M3 vs. M5b	3.44	2	.18
M6a: M3 + Δ SV + Δ SS	137.48	M3 vs. M6a	0.73	2	.69
M6b: M6a + Δ SV_Construction + Δ SS_Rule-Based + Δ Paragraph_Level	130.19	M3 vs. M6b	8.02	5	.15
M6c: M3 + Δ SV + Δ SV_Construction	131.64	M3 vs. M6c	6.57	2	.04
M6d: M6c + (Δ SV \times Timepoint) + (Δ SV_Construction \times Timepoint)	129.20	M6c vs. M6d	2.44	2	.29
M7a: M6c + Δ MO + (Δ MO \times Timepoint) + Δ NI + (Δ NI \times Timepoint)	122.26	M6c vs. M7a	9.39	4	.05
M7b: M6c + Δ MO + Δ NI + (Δ NI \times Timepoint)	122.25	M7a vs. M7b	1.83	1	.18

Note. *df* = degrees of freedom; MO = teacher modeling; NI = no individual activities; SN = structure number; SS = structure summarization; SV = structure visualization. The model fit for reduced M6c is not different from the full factorial M6b, $\Delta x^2(3) = 1.46, p = .69$; similarly, the model fit for reduced M7b is not different from the full M7a, $\Delta x^2(1) = 1.83, p = .18$. Parameter estimates are based on M7b.

Parameter Estimates of the Final Model for Comprehension Questions

Model result	Estimate	Standard error	<i>p</i>
Intercept	0.98	0.30	.001
Delayed posttest	-1.21	0.37	.001
Structure-based visualization (overall)	-0.29	0.15	.06
Active structure mapping	0.39	0.15	.009
Teacher modeling	-0.09	0.10	.38
No individual activities: Posttest	0.45	0.23	.06
No individual activities: Delayed posttest	-1.04	0.39	.007

Recall

Overview of Stepwise Model Fitting for Recall

Model (M)	-2 log likelihood ratio	Comparison	Δx^2	Δdf	<i>p</i>
M0	133.03				
M1: M0 + Δ Publication	128.03	M0 vs. M1	5.00	1	.025
M2: M0 + Δ Competing_Intervention	132.59	M0 vs. M2	0.44	1	.51
M3: M0 + Δ Timepoint	128.18	M0 vs. M3	4.84	1	.028
M4a: M3 + Δ Genre	128.08	M3 vs. M4a	0.11	1	.74
M4b: M3 + Δ Genre + (Δ Genre \times Timepoint)	128.16	M3 vs. M4b	1.01	2	.90
M5a: M3 + Δ SN	128.18	M3 vs. M5a	0.003	1	.96
M5b: M3 + Δ SN + (Δ SN \times Timepoint)	128.77	M3 vs. M5b	0.02	2	.99
M6a: M3 + Δ SV + Δ SS	128.10	M3 vs. M6a	0.08	2	.96
M6b: M6a + Δ SV_Construction + Δ SS_Rule-Based + Δ Paragraph_Level	112.69	M3 vs. M6b	15.49	5	.002
M6c: M6b + Timepoint(Δ SV_Construction + Δ SS_Rule-Based + Δ Paragraph_Level)	109.66	M6b vs. M6c	3.03	4	.55
M7: M6b + Δ MO + (Δ MO \times Timepoint) + Δ NI	110.18	M6b vs. M7	2.51	3	.47

Note. *df* = degrees of freedom; MO = teacher modeling; NI = no individual activities; SN = structure number; SS = structure summarization; SV = structure visualization. At model M7, none of the interventions without individual activities administered a delayed posttest. Parameter estimates are based on M6b.

Parameter Estimates of the Final Model for Recall

Model result	Estimate	Standard error	<i>p</i>
Intercept	1.03	0.39	.009
Delayed posttest	-0.18	0.11	.09
Paragraph-level structure instruction	0.57	0.29	.03
Structure-based visualization (overall)	-0.44	0.18	.02
Active structure mapping	0.51	0.17	.002
Structure-based summarization (overall)	-0.24	0.17	.14
Rule-based summarization technique	0.34	0.12	.004

Summarization

Overview of Stepwise Model Fitting for Summarization

Model (M)	-2 log likelihood ratio	Comparison	Δx^2	Δdf	<i>p</i>
M0	112.64				
M1: M0 + Δ Publication	108.69	M0 vs. M1	3.95	1	.05
M2: M0 + Δ Competing_Intervention	110.49	M0 vs. M2	2.16	1	.14
M3: M0 + Δ Timepoint	102.01	M0 vs. M3	9.73	1	.002

(continued)

Overview of Stepwise Model Fitting for Summarization (continued)

Model (M)	-2 log likelihood ratio	Comparison	Δx^2	Δdf	<i>p</i>
M4a: M3 + ΔSN	102.70	M3 vs. M4a	0.21	1	.65
M4b: M3 + ΔSN + ($\Delta SN \times$ Timepoint)	102.71	M3 vs. M4b	0.21	2	.90
M5a: M3 + ΔSV + ΔSS	97.13	M3 vs. M5a	5.78	2	.056
M5b: M5a + ΔSV _Construction + ΔSS _Rule-Based + Δ Paragraph_Level	81.31	M3 vs. M5b	15.49	5	.002
M5c: M3 + ΔSS + ΔSS _Rule-Based + Δ Paragraph_Level	82.10	M5b vs. M5c	0.79	2	.67
M5d: M5c + (Timepoint \times M5c)	80.78	M5c vs. M5d	1.32	3	.72
M6: M5c + ΔMO + ($\Delta MO \times$ Timepoint) + ΔNI	79.15	M5c vs. M6	9.39	3	.40

Note. *df* = degrees of freedom; MO = teacher modeling; NI = no individual activities; SN = structure number; SS = structure summarization; SV = structure visualization. The model fit for reduced M5c is not different from the full M5b, $\Delta x^2(2) = 0.79$, *p* = .67. For M6, there were no interventions with a delayed posttest without individual practice during the intervention. Parameter estimates are based on M5c.

Parameter Estimates of the Final Model for Summarization

Model result	Estimate	Standard error	<i>p</i>
Intercept	0.22	0.40	.58
Delayed posttest	-0.47	0.16	.004
Paragraph-level structure instruction	0.91	0.22	<.001
Structure-based summarization (overall)	-0.14	0.26	.61
Rule-based summarization technique	0.64	0.21	.005

Text Structure Knowledge

Overview of Stepwise Model Fitting for Text Structure Knowledge

Model (M)	-2 log likelihood ratio	Comparison	Δx^2	Δdf	<i>p</i>
M0	12.16				
M1: M0 + Δ Publication	12.39	M0 vs. M1	0.23	1	.63
M2: M0 + Δ Competing_Intervention	13.30	M0 vs. M2	1.14	1	.29
M3: M0 + ΔSN	16.88	M0 vs. M3	4.64	1	.03

Note. *df* = degrees of freedom; SN = structure number. Because of a low number of observations, the moderating impact of other content-related and instructional variables was not analyzed. Parameter estimates are based on M3.

Parameter Estimates of the Final Model for Text Structure Knowledge

Model result	Estimate	Standard error	<i>p</i>
Intercept	0.57	0.11	<.001
Number of different structures taught	-0.06	0.02	.02