

Differences in text processing by low- and high-comprehending beginning readers of expository and narrative texts: Evidence from eye movements



Astrid Kraal^{a,*}, Paul W. van den Broek^a, Arnout W. Koorneef^a, Lesya Y. Ganushchak^b, Nadira Saab^c

^a Department of Education and Child Studies, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, the Netherlands

^b Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands

^c ICLON, Leiden University Graduate School of Teaching, Kolffpad 1, 2333 BN Leiden, the Netherlands

ARTICLE INFO

Keywords:

Reading comprehension
Eye movements
On-line text processing
Low- and high-comprehending beginning readers
Narrative and expository texts

ABSTRACT

The present study investigated on-line text processing of second-grade low- and high-comprehending readers by recording their eye movements as they read expository and narrative texts. For narrative texts, the reading patterns of low- and high-comprehending readers revealed robust differences consistent with prior findings for good versus struggling readers (e.g., longer first- and second-pass reading times for low-comprehending readers). For expository texts, however, the differences in the reading patterns of low- and high-comprehending readers were attenuated. These results suggest that low-comprehending readers adopt a suboptimal processing approach for expository texts: relative to their processing approach for narrative texts, they either do not adjust their reading strategy or they adjust towards a more cursory strategy. Both processing approaches are suboptimal because expository texts tend to demand more, rather than less, cognitive effort of the reader than narrative texts. We discuss implications for (reading) education.

1. Introduction

Reading comprehension is one of the most important skills children need to acquire in school, as much academic knowledge is transferred through texts. However, individual differences in comprehension skills are considerable, and many children fail to comprehend the texts they have to read in school (e.g., Kuhlemeier et al., 2014; National Center for Education Statistics, 2011). Whereas in the early grades of elementary school children primarily come into contact with narrative texts, from third grade onward expository texts become more dominant. Expository texts are an important learning tool aimed at transmitting knowledge about a particular theme or subject. However, the processing of expository texts, relative to that of narrative texts, poses a challenge for young readers for several reasons. Young readers often are unfamiliar with the conventions, the global structure, and the reading goals for expository texts (Cotterill, Goldman, & Saul, 1998; Duke, 2000; Kaakinen, Hyvri, & Keenan, 2003; Lorch, 2017; Meyer, 1975, 1985; Meyer & Freedle, 1984; Meyer & Ray, 2011; Oakhill, Cain, & Elbro, 2014; Williams et al., 2007; Williams, Hall, & Lauer, 2004). Furthermore, because expository texts often contain unfamiliar words and concepts (i.e., young readers are likely to have gaps in their prior

knowledge), they also require strong inferential skills (Kaakinen et al., 2003; Lorch, 2017; Oakhill et al., 2014). As a result, most children understand narrative texts better than expository texts (Best, Floyd, & McNamara, 2008) with expository texts being particularly difficult to comprehend for struggling readers (Williams et al., 2004).

To have an understanding of where exactly these text comprehension problems occur and to be able to alleviate them, it is necessary to gain insight into the processes in which readers engage during reading. Tracking the eye movements of readers during reading can reveal readers' on-line text-processing approaches (Boland, 2004; Juhasz & Pollatsek, 2011; Rayner, 1998). In the current study, we tracked the eye movements of young low- and high-comprehending readers to investigate how they process narrative and expository texts, and how they may differ in terms of on-line text processing.

1.1. Individual differences in reading and eye movements

It generally is presumed that readers employ strategies to facilitate comprehension, either consciously or unconsciously (Pressley & Afflerbach, 1995). Sources of individual differences in reading comprehension include general cognitive factors (e.g., background

* Corresponding author at: Department of Research and Innovation, CED-Group, PO Box 8639, 3009 AP Rotterdam, the Netherlands

E-mail addresses: a.kraal@fsw.leidenuniv.nl, a.kraal@cedgroep.nl (A. Kraal), p.vandenbroek@fsw.leidenuniv.nl (P.W. van den Broek), a.w.koorneef@fsw.leidenuniv.nl (A.W. Koorneef), ganushchak@essb.eur.nl (L.Y. Ganushchak), n.saab@iclon.leidenuniv.nl (N. Saab).

<https://doi.org/10.1016/j.lindif.2019.101752>

Received 7 February 2018; Received in revised form 26 June 2019; Accepted 1 July 2019

1041-6080/© 2019 Elsevier Inc. All rights reserved.

knowledge, working-memory capacity, vocabulary, reasoning skills), comprehension factors (e.g., standards of coherence, inference-making skills), and text-specific factors (e.g., knowledge about text structures). For reviews see Cain & Oakhill, 2007; Helder, Van den Broek, Van Leijenhof, & Beker, 2013; Van den Broek & Espin, 2012). A first common finding in research on comprehension processes of successful readers is that they tend to be active, strategic readers (Duke & Pearson, 2002; Pressley & Wharton-McDonald, 1997; Rapp, Van den Broek, McMaster, Kendeou, & Espin, 2007; Van der Schoot, Reijntjes, & Van Lieshout, 2012). They read texts with a clear goal (Duke & Pearson, 2002; Pressley & Wharton-McDonald, 1997), read different kinds of texts differently (Duke & Pearson, 2002; Schroeder, 2011; Zaburky & Ratner, 1992), and regulate their reading strategies depending on the text and the task (Rapp et al., 2007; Schroeder, 2011). A second common finding is that successful readers make inferences while reading. They relate important ideas in the text to one another, and to their prior knowledge (Duke & Pearson, 2002; Pearson, Roehler, Dole, & Duffy, 1992; Pressley & Wharton-McDonald, 1997; Rapp et al., 2007). A third common finding is that successful readers tend to be highly metacognitive. They monitor their reading and, when they encounter a comprehension problem, try to solve it (Pearson et al., 1992; Pressley & Wharton-McDonald, 1997). Sometimes they jump ahead in the text, or they look back when they are confused (Pressley & Wharton-McDonald, 1997). Further, they cope with discrepancies and gaps and attempt to define the meanings of words and concepts they do not know (Duke & Pearson, 2002). Thus, successful readers set themselves high standards of coherence (Van den Broek, Bohn-Gettler, Kendeou, Carlson, & White, 2011). They have a drive to understand the texts they read and put cognitive effort into it. By doing all this, successful readers construct a coherent mental representation of the meaning of a text (e.g., Van den Broek, White, Kendeou, & Carlson, 2009).

Readers that struggle with comprehension of the texts they read differ from their successful peers in several aspects. One difference concerns the goals they set themselves (Oakhill & Cain, 2007). Readers that are unsuccessful in comprehending texts regard reading more as a matter of word decoding than a matter of sense-making (Cain, 1999; Yuill & Oakhill, 1991), and they are less likely than successful readers to adapt their reading style to different reading tasks (Cain, 1999). Furthermore, unsuccessful readers may have difficulty with generating the correct inferences when needed, with identifying particular relations in the text, or with integrating information from the text with background knowledge (Cain, Oakhill, Barnes, & Bryant, 2001; Denton et al., 2015; Oakhill & Cain, 2007; Perfetti, Landi, & Oakhill, 2005; Yuill & Oakhill, 1991). In addition, unsuccessful readers frequently are less proficient than successful readers in comprehension monitoring (Denton et al., 2015; Oakhill & Cain, 2007; Oakhill, Hartt, & Samols, 2005; Perfetti et al., 2005). Finally, whereas successful readers use their knowledge of text structure to relate ideas in the text and to activate relevant text schemas, unsuccessful readers are less sensitive to story structure (Oakhill & Cain, 2007, 2012). Consequently, unsuccessful readers are less conscious of the features of texts that could help them construct a coherent mental representation of the text's meaning (Oakhill & Cain, 2012; Perfetti et al., 2005).

Differences in reading skills may be reflected in differences in eye-movement patterns. In eye-tracking experiments it is common to relate eye-movement measures to cognitive processes (Boston, Hale, Kliegl, Patil, & Vasishth, 2008; Clifton, Staub, & Rayner, 2007; Juhasz & Pollatsek, 2011; Kuperman & Van Dyke, 2011; for an overview of commonly used eye-movement measures and the cognitive processes they are assumed to represent, see e.g., Boston et al., 2008). Measures of initial or early processing (first-pass measures) are assumed to represent lower-level processes and skills such as decoding of orthographic information and word identification (Boland, 2004; Boston et al., 2008; Juhasz & Pollatsek, 2011; Kuperman & Van Dyke, 2011). These measures include, for instance, *first fixation duration* (the duration of the first fixation on a word during first pass; Juhasz & Pollatsek,

2011), *first gaze duration* (also *gaze duration* or *first-pass reading time*: the sum of all fixation durations for a word before the reader moves on or looks back in the text; Boston et al., 2008; Juhasz & Pollatsek, 2011), and *first-pass word skipping probability* (or *skipping rates*: the probability that a word is skipped during first pass; Juhasz & Pollatsek, 2011). Later processing measures are assumed to represent higher-level processes and skills involved in integrative text processing (Boland, 2004; Boston et al., 2008; Juhasz & Pollatsek, 2011; Kuperman & Van Dyke, 2011). Later processing measures include, for instance, *regression rates* (or *regression probability*: the percentage of regressions into a target word or out of a target word; Boston et al., 2008; Juhasz & Pollatsek, 2011), *second-pass fixation duration* (or *re-reading time*: the amount of time that is spent on re-reading a target word after first-pass reading; Boston et al., 2008; Juhasz & Pollatsek, 2011), *total fixation duration* (the sum of all fixation durations on a word; Juhasz & Pollatsek, 2011), and *right-bounded duration* (or *right-bounded reading time*: the sum of all fixation durations for a word during first pass before the reader moves on progressively; Boston et al., 2008; Traxler, Bybee, & Pickering, 1997).

The distinction between initial versus later processing measures has yielded important insights into readers' text processing. For instance, with regard to initial processing measures reflecting lower-level processes and skills, prior research has shown that both beginning readers and poor readers make more fixations, longer fixations, shorter saccades, and more regressions when reading texts than proficient readers do (Rayner, 1998). However, the increased regression rates of struggling readers do not imply that regressive eye movements should be avoided. In fact, they can fulfil an important purpose, namely to repair a failure in text comprehension (Rayner, Schotter, Masson, Potter, & Treiman, 2016). With regard to later processing measures, which reflect higher-level processes and skills that lead to the construction of a mental model of the text, prior research has shown that the effective comprehension-monitoring skills of good comprehenders may result in these readers making more regressive eye movements than do poor comprehenders (Ehrlich, Remond, & Tardieu, 1999; Rayner et al., 2016; Van der Schoot et al., 2012). By investigating initial and later processing measures one can obtain insight into both lower-level processes and higher-level cognitive processes that readers engage in during reading and, possibly, into differences between subgroups of low- and high-comprehending readers. Although one-to-one mapping between specific eye movements and specific cognitive processes is difficult (Boland, 2004; Kuperman & Van Dyke, 2011), the general mapping of early and late measures is widely accepted.

1.2. Differences in reading and eye movements as a function of text genre

Differences in reading processes and eye-movement patterns depend not only on the skills of the reader but also on properties of the text (e.g., see Huestegge & Bocianski, 2010; Juhasz & Pollatsek, 2011; Oakland & Lane, 2004). Text genre may play an important role in readers' text-processing styles (McDaniel & Einstein, 1989; Oakhill & Cain, 2012; Van den Broek & Kendeou, 2017). As pointed out earlier, expository texts are more challenging than narrative texts to many children (Best et al., 2008). Children may need to make extensive use of comprehension processes such as elaborative inference making (using their prior knowledge to make sense of the information in the text) and comprehension monitoring to meet the challenge of reading these texts. However, their lack of familiarity with the content and structures of expository texts may hamper them in doing this successfully (Cotter et al., 1998). In fact, with regard to adolescents Denton et al. (2015) have shown that poor comprehenders have considerable deficits in inference making, particularly in expository text, and in comprehension monitoring in both expository and narrative texts.

To our knowledge, no previous eye-tracking studies have investigated young readers' processing of narrative texts relative to expository texts. However, there is evidence from previous eye-tracking studies that readers process easy texts and difficult texts differently. For

difficult texts, readers have longer reading times and longer fixation durations, make more fixations, and skip fewer words (Rayner, 1986). Furthermore, for difficult texts readers look back more often and spend more time looking back to previous parts than for easy texts (Leeuw, Segers, & Verhoeven, 2016). Based on these findings, it is plausible that more cognitive effort is exerted in processing expository texts than in processing narrative texts, and that this leads to a more effortful or intensive reading style. However, it is not sure whether low-comprehending readers are capable of engaging in a more effortful reading style of this kind. Denton et al. (2015) indicate that whereas good comprehenders exert more effort to process difficult texts, poor comprehenders fail to engage in more effortful processing.

Investigation of individual differences in text processing and in the flexibility to adapt to text genre (either automatically or strategically) can provide important insights into comprehension differences. Previous studies on children's eye movements when reading have focused on one text genre (e.g., de Leeuw, Segers, & Verhoeven, 2016; Kaakinen, Lehtola, & Paattilammi, 2015; Olson, Kliegl, Davidson, & Foltz, 1985; Rayner, 1986; Van der Schoot et al., 2012), or on specific word or text manipulations (e.g., Blythe et al., 2006; Blythe, HÃ¶rri, Bertam, Liversedge, & HyÃ¶nÃ¤, 2011; Blythe, Liversedge, Joseph, White, & Rayner, 2009; Huestegge & Bocianski, 2010; Huestegge, Radach, Corbic, & Huestegge, 2009; Joseph, Bremner, Liversedge, & Nation, 2015; Joseph, Liversedge, Blythe, White, & Rayner, 2009; Schroeder, 2011). With regard to lower-level processes and skills involved in reading, it is likely that poor readers differ from proficient readers in their eye movements (e.g., see Rayner, 1998). However, with regard to higher-level processes and skills in reading, the picture of how young low- and high-comprehending readers differ in their eye movements is less complete. Neither is it clear whether their text processing as reflected by their eye-movement patterns differs for different types of texts (expository vs. narrative texts). It is important to obtain a clear understanding of the processes involved in young low- and high-comprehending readers' on-line reading comprehension, especially for different types of texts, as this may shed light on aspects of reading that hinder text comprehension for young unsuccessful readers. Tracking children's eye movements can contribute to obtaining insights into readers' on-line text processing.

1.3. The present study

This study is part of a larger study consisting of several experiments in which various aspects of reading comprehension were investigated. In a prior study we examined the on-line comprehension processes and strategy use of young low- and high-comprehending readers by using a think-aloud protocol (Kraal et al., 2018). The purpose of the present study is to gain insight into the on-line text processing of young low- and high-comprehending readers by tracking their eye movements as they read expository and narrative texts for comprehension. We hypothesized that because young readers have greater difficulty processing expository texts than narrative texts (Best et al., 2008), they need to adapt their reading approach to obtain a proper understanding of expository texts (cf. Leeuw et al., 2016; Rayner, 1986). In addition, we hypothesized that readers with high comprehension skills will adjust their reading approach more easily to the text than readers with poor comprehension skills (Duke & Pearson, 2002; Rapp et al., 2007) because high-comprehending readers generally have better comprehension-monitoring skills (Cain & Oakhill, 2007; Perfetti et al., 2005; Van der Schoot et al., 2012), better strategic skills (Duke & Pearson, 2002; Pressley & Wharton-McDonald, 1997; Rapp et al., 2007; Van der Schoot et al., 2012), greater structural awareness (Ray & Meyer, 2011), and a more flexible reading approach (Duke & Pearson, 2002; Rapp et al., 2007; Schroeder, 2011; Zabrocky & Ratner, 1992). Based on these hypotheses, we predicted an interaction effect between reading proficiency (high vs. low) and text genre (narrative vs. expository) for the eye-movement measures that reflect higher-level integrative text

processing and comprehension (e.g., regression rates, re-reading times; Boland, 2004; Boston et al., 2008; Juhasz & Pollatsek, 2011; Kuperman & Van Dyke, 2011).

To explore these issues, we asked second-grade low- and high-comprehending pupils to read expository and narrative texts from a computer screen in an eye-tracking experiment. They were all proficient technical readers with a normally developed intellect. After reading, the pupils answered literal and inferential comprehension questions to determine the quality of their mental text representation (Cain & Oakhill, 1999). In the context of the larger study they completed a test battery measuring general cognitive and language-related skills to identify possible factors contributing to comprehension and processing differences (Kraal et al., 2018).

2. Method

2.1. Participants

Eighty second-grade pupils (47 girls; mean age 7:8, range 6:4â€“8:9) from 19 classes of 9 elementary schools in the Netherlands participated in this study. They were selected from a larger screening sample ($N = 385$) (Kraal et al., 2018). Following Cain and Oakhill (2007) and Nation (2005), we defined our low-comprehending readers as readers who have difficulty comprehending connected text despite having age-appropriate technical reading skills. Selection criteria were as follows: Average or above-average score on a standardized test for word-reading ability (*Drie-Minuten-Test* [Three-Minutes-Test], Cito, 2009); average or above-average IQ-score on a non-verbal intelligence test (*Raven's progressive matrices*, Raven, Raven, & Court, 1998); either a score above the 75th percentile or below the 50th percentile score on a standardized test for reading comprehension (*LOVS Begrijpend Lezen Groep 3* [Reading Comprehension Grade 1], Cito, 2006); no diagnosed behavioral and/or attentional problems. On the basis of their score on the standardized test for reading comprehension, the selected children were assigned to a group of high-comprehending readers ($N = 53$) or to a group of low-comprehending readers ($N = 27$). The 50th percentile was applied as a threshold for classifying low-comprehending readers because a score below this percentile indicates a score below the national average. Moreover, we ensured that there was a clear gap in comprehension skills between the low- and high-comprehending readers as we only included high-comprehending pupils that obtained scores above the 75th percentile.

The data of three low-comprehending participants were removed from the dataset due to tracker-loss in the eye-tracking part of the experiment. Therefore, reported analyses relate to 77 participants (high-comprehending, $N = 53$; low-comprehending, $N = 24$). Before testing, participants' parents or guardians provided written informed consent. After testing, the participants received an eraser and their teachers received a book token ($\text{â} = -20$).

2.2. Measures and materials

2.2.1. Stimuli of eye-tracking experiment

In the eye-tracking session, the children read one narrative practice text, two expository texts, and two narrative texts, all of which were especially designed for the current study. The texts were presented in the following order: Narrative practice text about a boy who helps his uncle milking cows on a farm; expository text about the human skeleton; narrative text about two brothers who encounter a problem with their sister's iPad; expository text about lions; narrative text about children playing hide-and-seek in the schoolyard. The texts are provided in Appendix A. The texts were matched on length and readability, using an algorithm that calculates text difficulty at the level of conceptual readability (P-CLIB version 3.0, Evers, 2008). See Table 1 for details on text features of the experimental texts.

Table 1
Text features of narrative and expository experimental texts based on P-CLIB.

	Narrative texts		Expository texts	
	iPad	Hide-and-seek	Human skeleton	Lions
No. of sentences	19	19	19	19
No. of words/text	127	116	123	122
No. of types/text ^a	85	72	64	74
No. of letters/text	504	496	524	525
Average no. of words/sentence	6.7	6.1	6.4	6.4
Average no. letters/word	4.0	4.3	4.3	4.3
Syllables -, 1, 2, 3, N ^b	0, 87, 33, 7, 0	0, 83, 24, 8, 1	0, 76, 42, 5, 0	0, 80, 35, 5, 2
Percentage of frequent words	86.6	91.4	72.4	78.7
Type-token ratio	0.67	0.62	0.52	0.61
CLIB (grade level)	CLIB-4 ^c	CLIB-4	CLIB-5 ^c	CLIB-5

^a No. of types/text represents the number of distinct words in a text.

^b Syllables -, 1, 2, 3, N represent the number of words for which no syllables are counted (i.e., abbreviations and numbers not written as text), the number of one-syllable words, the number of two-syllable words, the number of three-syllable words and the number of words with more than three syllables, respectively.

^c CLIB-4 is the equivalent of the text-difficulty level for Grade 2 and CLIB-5 for Grade 3.

2.2.1.1. Eye movement measures. Full text reading times and Full text fixation counts were computed as coarse measures of processing strategies of different types of readers (low- vs. high-comprehending readers) for different types of texts (narratives vs. expository texts). In addition, to detect more subtle differences for these two factors, we computed several finer-grained reading measures for each individual word of a text (see e.g., Juhász & Pollatsek, 2011): *First fixation duration* (the duration of the first fixation on a word during first pass), *First gaze duration* (the sum of all fixation durations for a word during first pass before the reader either moves on or looks back in the text), *Right-bounded duration* (the sum of all fixation durations for a word during first pass before the reader moves on progressively), *Total fixation duration* (the sum of all fixation durations for a word), *Second-pass fixation duration* (the sum of all second-pass fixation durations for a word, i.e., computed by subtracting the right-bounded duration from the total fixation duration), *First-pass regression probability* (based on the ratio of progressive to regressive saccades during first pass), *First-pass word skipping probability* (the probability that a word is skipped during first pass) and *Saccade amplitude* (the amplitude of progressive first-pass saccades in degree of visual angle).

2.2.1.2. Comprehension questions. Three types of questions were included to assess text comprehension after reading (each reader answered six questions in total after reading each text): 1) two questions eliciting literal information; 2) three text-based questions requiring a text-connecting inference between two sentences in the text, and 3) one knowledge-based question requiring a gap-filling inference between the information in the text and the reader's background knowledge (Cain & Oakhill, 1999). The answers were scored as correct or incorrect. The comprehension questions are provided in Appendix A.

2.2.2. Test battery¹

2.2.2.1. Non-verbal intelligence (selection criterion). To assess non-verbal intelligence and abstract reasoning, Raven's Standard Progressive Matrices (Raven et al., 1998) were used. Participants completed the

¹ The test battery was administered as part of a larger study (Kraal et al., 2018). Here, only tests relevant to the current study are described.

test at their own pace in the classroom, marking their answers on an answer sheet. The maximum time on task was 30 min. Completion was not required. Raw scores are reported. The scores of the participants ranged from 15 to 49, reflecting average to high non-verbal intelligence (range of possible scores: 0–60).

2.2.2.2. Word-reading ability (selection criterion). To assess word-decoding skills, a standardized test was used (*Drie-Minuten-Toets* [Three-Minutes-Test], Cito, 2009). The test consists of three lists of words of increasing difficulty. Children read aloud as many words as they can from each list for one minute, with emphasis on both speed and accuracy. The reliability of the DMT is good, $\hat{I} \pm \hat{\sigma} > \hat{\sigma} 0.97$ (Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010). In all participating schools, the children took the test at the end of first grade. Skill scores are reported. The scores of the participants ranged from 34 to 85, reflecting above average to very high word-reading ability (range of possible scores: 0–148).

2.2.2.3. Reading comprehension (selection and classification criterion). To assess reading-comprehension skills, a Dutch standardized reading-comprehension test (*LOVS Begrijpend Lezen Groep 3* [Reading Comprehension Grade 1], Cito, 2006) was used. The reliability of this test is good, $MAcc > 0.93$ (Feenstra, Kamphuis, Kleintjes, & Krom, 2010). The test consists of three modules: An initial module for all children, an easier follow-up module for struggling comprehenders, and a more difficult follow-up module for average and good comprehenders. The modules contain different types of items: Shuffled stories, fill-in assignments, and questions about the text. The initial module has four example items and 25 multiple-choice items. The follow-up modules consist of 25 multiple-choice items each. The children took the test at the end of first grade in six of the nine schools. At three schools, we administered the initial module of the test instead of the school. The time on task was 30 min. Skill scores are reported. The scores of the participants ranged from 30 to 34, reflecting very low to very high reading-comprehension ability (range of possible scores for the easier follow-up module: 87 to 48; range of possible scores for the more difficult follow-up module: 81 to 60).

2.2.2.4. Listening comprehension. To measure listening-comprehension skills a standardized listening-comprehension test was used (*LOVS Begrijpend Luisteren Groep 3* [Listening Comprehension Grade 1], Cito, 2011). The reliability of this test is adequate, $MAcc > 0.72$ (Van Berkel et al., 2013). The test includes two parts consisting of 23 multiple-choice questions each. In each part, the children listen to one- to four-sentence stories and answer a question by choosing the correct picture from three available pictures. Administration of the test took 25 min per part. There was a one-week interval between the administration of part 1 and 2 to avoid fatigue and loss of concentration. Skill scores are reported. The scores of the participants ranged from 34 to 83, reflecting the full range of listening-comprehension ability (range of possible scores: 0–113).

2.2.2.5. Vocabulary knowledge. To assess receptive vocabulary in Dutch, the Peabody Picture Vocabulary Test-III-NL (Schlichting, 2005) was used as a standardized measure. The test includes sixty words that differ in difficulty. Each word is presented together with four pictures. The child has to indicate which picture matches the target word. Administration of the task took 25 min. Raw scores are reported. The scores of the participants ranged from 22 to 52 (range of possible scores: 0–60).

2.2.2.6. Verbal working memory. To assess verbal working-memory capacity, a translated and adapted version of the Sentence Span measure (Swanson, Cochran, & Ewers, 1989) was used. The original test comprises four levels with two sets of unrelated sentences, with levels increasing in difficulty: The lowest level contains two sets of two

sentences, the highest level contains two sets of five sentences. We translated the test. In a pilot test, the test proved to be too difficult for many children. Therefore, we added an easier level comprising two sets of one sentence of seven to ten words in length (as in the original test). The test was administered individually. Scores for the number of correctly remembered words plus correctly answered comprehension questions are reported. This scoring method has good internal consistency of 0.79 (Conway et al., 2005). The scores of the participants ranged from 1 to 16 (range of possible scores: 0–30).

2.2.2.7. Inference making. To assess inference-making skills a translated version of the Inference and Integration Task (Cain & Oakhill, 1999) was used. The test contains one practice story and three test stories with six comprehension questions each: Two questions eliciting literal information, two questions requiring a text-based inference, and two questions requiring a knowledge-based inference. Raw scores are reported. The scores of the participants ranged from 7.5 to 17 (range of possible scores: 0–18).

2.3. Procedure

All procedures were approved by the Leiden University Institute of Education and Child Studies ethics committee (project number ECPW-2014/073) and conducted in accordance with the Declaration of Helsinki. Participants were tested in two phases, in five sessions in total, see Table 2. In the first three sessions of phase 1, the larger screening sample (N = 385) was tested on general cognitive and language-related proficiencies. After this first phase, the selection criteria (see Section 2.1 Participants) were applied; resulting in the selection of 80 children. In the second phase, the selected children were further tested in two individual sessions, which took place within a one-week-interval (by this stage the children had spent six months in Grade 2). In the first individual session (30 min) inference-making skills and verbal working-memory capacity were assessed. In the second individual test session (45–60 min), the eye-tracking experiment was conducted. An Eyelink 1000 setup (SR Research) was used for the experiment, sampling at a rate of 500 Hz. Children were tested at their schools in a quiet and dim room. Children were first familiarized with the eye tracker and then instructed to read the texts for comprehension and to answer comprehension questions after reading each text.

All texts were displayed 60 cm from the participants' eyes, on a computer screen (Dell, P1913sb, 19 in, 60 Hz). The texts were presented in their entirety (one sentence per line) without title, in black Tahoma font 15 pt. size on a light grey background, with double spacing between the lines. The children rested their head on a chin-rest to prevent them from moving their head during reading. Before the children read each text, their right eye was calibrated using nine black dots on a white background. Each trial began with a fixation point presented about two-character spaces to the left of the first character of the upcoming sentence. The children's eyes were calibrated five times: Before they read the practice text and before they read each experimental text. The task was self-paced. The children read the texts silently for both theoretical and practical reasons. Reading aloud may influence

Table 2
Testing phases, sessions and tests.

	Phase 1 Group-administered sessions	Phase 2 Individual sessions
Session 1	-Listening comprehension part 1	-Inference making
		-Verbal working memory
Session 2	-Listening comprehension part 2	-Eye-tracking experiment
	-Vocabulary knowledge	
Session 3	-Non-verbal intelligence	
	-Reading comprehension initial module ^a	

^a If not administered previously by the school.

the reader's processing approach. The processes involved in reading aloud (e.g., speech articulation, articulation monitoring, eye-voice coordination) may impose extra demands on the reader's working memory and prevent readers from making regressions (Vorstius, Radach, & Lonigan, 2014). In addition, the fact that the reader must articulate every word possibly leads to more emphasis on sublexical word properties (Huestegge, 2010). A practical reason is that reading aloud is not a good option with a chin-rest, used to prevent participants' head movements and allow accurate eye-movements measurements. After completion of each text, the text was removed from the screen and the child moved his/her head from the chin-rest and answered the six comprehension questions orally. The questions were posed by the experimenter.

3. Results

3.1. Test battery

On the tests for inference making and verbal working memory data of one low-comprehending participant were missing. On the test for inference making data of one high-comprehending participant were missing. On the test for verbal working memory data of two high-comprehending participants were missing. Fourteen participants (2 low- and 12 high-comprehending readers) only made the initial part of the test for reading comprehension.

A MANOVA was performed, with Reading Proficiency as a between-participants factor and the scores for the cognitive and language-related tasks in the test battery as dependent measures. The MANOVA showed that the group of high-comprehending readers performed better on the tasks than the group of low-comprehending readers ($F(7,52) = 25.58, p < .001$). Univariate ANOVAs revealed that this difference between groups was reliable for all tasks ($p < .01$; see Table 3 for descriptive statistics of the dependent measures).

3.2. Comprehension questions

Table 4 reports the mean values of the accuracy scores of the comprehension questions as a function of the factors Text Genre (expository vs. narrative), Question Type (literal, text-based, and knowledge-based questions), and Reading Proficiency (high vs. low).

A mixed-effects logistic regression model was fitted on the data with the R package LME4 (R-version 3.3.3; LME4 version 1.1–12). First, a model was fitted that included the fixed factors Text Genre, Reading Proficiency, and Question Type, as well as the interactions between these factors. Participants and questions were included as crossed random effects (Baayen, Davidson, & Bates, 2008; cf. Barr, Levy, Scheepers, & Tily, 2013, and Bates, Kliegl, Vasishth, & Baayen, 2015, for discussion). Because this model did not converge, models were fitted for each type of question separately (i.e., these models included the fixed factors Text Genre and Reading Proficiency, their interaction, and

Table 3
Mean scores (and standard deviations) of high- and low-comprehending readers on tasks in the test battery.

Measure	Reading proficiency	
	High	Low
	<i>M (SD)</i>	<i>M (SD)</i>
Non-verbal intelligence	36.02 (6.24)	27.79 (7.31)
Word-reading ability	54.68 (12.53)	46.88 (9.49)
Reading comprehension	18.78 (7.61)	8.27 (6.61)
Listening comprehension	61.08 (8.18)	49.50 (9.61)
Vocabulary knowledge	42.08 (6.13)	33.29 (6.54)
Verbal working memory	5.80 (3.35)	2.74 (1.18)
Inference making	14.70 (1.45)	12.37 (2.38)

Table 4
Mean accuracy scores of high- and low-comprehending readers on comprehension questions for expository and narrative texts^a.

Question type	Text genre	Reading proficiency	
		High	Low
Literal	Expository	0.61	0.32
	Narrative	0.94	0.86
Text-based	Expository	0.57	0.26
	Narrative	0.77	0.52
Knowledge-based	Expository	0.71	0.48
	Narrative	0.83	0.32

^a Standard deviations (SD) do not apply to binary data and will only be reported for the continuous dependent variables.

random effects of participants and questions). Subsequently, Wald chi-square testing (Type II) χ^2 as implemented in the R package Car (version 2.1.4) was applied to select the most parsimonious model by removing non-significant predictors (i.e., $p > .05$, unless mentioned otherwise). For the final models obtained by this procedure, the relevant fixed-effects estimates and z-values will be reported. Statistical significance at approximately the 0.05 level is indicated by z-values of ≥ 1.96 or ≤ -1.96 (see e.g., Schotter, Tran, & Rayner, 2014). In the case of an interaction effect, fixed-effects estimates and the associated statistics for all relevant simple effects were obtained by adjusting the reference category of the models and fitting analogous models with the same structure of fixed factors and interactions.

3.2.1. Literal questions

The Wald chi-square tests revealed main effects of Text Genre ($\chi^2(1) = 12.5, p < .001$) and Reading Proficiency ($\chi^2(1) = 14.8, p < .001$). Accuracy scores of the literal questions were higher for narrative texts than for expository texts ($b = 3.24, SE = 0.89, z = 3.63$), and high-comprehending readers obtained higher scores than low-comprehending readers ($b = 1.62, SE = 0.41, z = 3.95$).

3.2.2. Text-based questions

The Wald chi-square tests revealed main effects of Text Genre ($\chi^2(1) = 7.42, p < .01$) and Reading Proficiency ($\chi^2(1) = 40.9, p < .001$). Accuracy scores of the text-based questions were higher for narrative texts than for expository texts ($b = 1.18, SE = 0.43, z = 2.72$), and high-comprehending readers obtained higher scores than did low-comprehending readers ($b = 1.43, SE = 0.22, z = 6.40$).

3.2.3. Knowledge-based questions

The Wald chi-square tests revealed a main effect of Reading Proficiency ($\chi^2(1) = 32.9, p < .001$) and an interaction effect between Text Genre and Reading Proficiency ($\chi^2(1) = 4.48, p < .05$). High-comprehending readers obtained higher accuracy scores than low-comprehending readers for both expository ($b = 1.19, SE = 0.41, z = 2.87$) and narrative texts ($b = 2.42, SE = 0.44, z = 5.54$). As can be seen in Table 4, the magnitude of this difference between low- and high-comprehending readers was smaller for expository texts.

3.3. Eye-tracking experiment

3.3.1. Eye-movement measures

Prior to all analyses, trials with full text reading times below 10s were removed from the dataset (< 1%). Furthermore, words that were not fixated by the reader were treated as missing data in the word-based

Table 5

Mean values (and standard deviations) of the eye-movement measures as a function of Reading Proficiency and Text Genre. First fixation duration and first gaze duration are considered to be ‘early’ measures of processing. Right-bounded duration, second-pass fixation duration, and regression probability reflect measures of ‘later’ processing. Word skipping probability and saccade amplitude are more difficult to categorize (they may reflect early measures of processing, but they may also be sensitive to strategic adjustments in reading).

Eye-movement measures	Text genre	Reading proficiency	
		High	Low
		<i>M (SD)</i>	<i>M (SD)</i>
Full text reading time ^a	Expository	50.9 (15.3)	54.6 (10.4)
	Narrative	49.4 (14.6)	56.3 (10.9)
Full text fixation count	Expository	181 (41)	180 (32)
	Narrative	175 (39)	185 (32)
First fixation duration ^b	Expository	285 (149)	310 (158)
	Narrative	284 (156)	307 (157)
First gaze duration ^b	Expository	384 (263)	427 (289)
	Narrative	368 (242)	407 (258)
Right-bounded duration ^b	Expository	459 (349)	494 (336)
	Narrative	431 (300)	474 (307)
Total fixation duration ^b	Expository	517 (398)	545 (391)
	Narrative	487 (352)	539 (369)
Second-pass fixation duration ^b	Expository	445 (425)	417 (288)
	Narrative	421 (309)	483 (377)
Saccade amplitude ^c	Expository	2.30 (1.88)	2.22 (2.09)
	Narrative	2.27 (1.86)	2.14 (2.02)
Regression probability	Expository	0.25	0.20
	Narrative	0.25	0.22
Word skipping probability	Expository	0.33	0.32
	Narrative	0.34	0.28

^a In seconds.

^b In milliseconds.

^c In degrees of visual angle.

reading time measures. Table 5 reports the mean values (and their SDs) of the eye-movement measures computed from the remaining data as a function of Text Genre and Reading Proficiency.

The analyses were conducted in the same way as described in Section 3.2. Mixed-effects linear regression models for the continuous reading time data with the response variable log-transformed to correct for right skewness and mixed-effects logistic regression models for the categorical dependent measures were fitted on the data. First, models were fitted for each dependent variable. Each model included the fixed factors Text Genre and Reading Proficiency, as well as the interaction of these factors. Participants and texts were included as crossed random effects. Subsequently, Wald chi-square testing was applied to select the most parsimonious model. The relevant fixed-effects estimates and the associated t-values (for the continuous dependent variables) and z-values (for the categorical dependent variables) will be reported.

3.3.1.1. Full text reading time. The Wald chi-square tests revealed a main effect of Reading Proficiency ($\chi^2(1) = 4.14, p < .05$) and an interaction effect between Text Genre and Reading Proficiency ($\chi^2(1) = 4.02, p < .05$). Further inspection of the interaction indicated that low-comprehending readers ($M = 56.32, SD = 10.92$) spent more time reading narrative texts than did high-comprehending readers ($M = 49.39, SD = 14.56$), ($b = 0.15, SE = 0.061, t = 2.43$). For expository texts, this contrast between low-comprehending readers ($M = 54.65, SD = 10.45$) and high-comprehending readers ($M = 50.89, SD = 15.31$) was attenuated ($b = 0.094, SE = 0.061, t = 1.54$). As illustrated in Fig. 1A, the effect of reading proficiency was attenuated in expository texts because high-comprehending readers had the tendency to slow down for expository

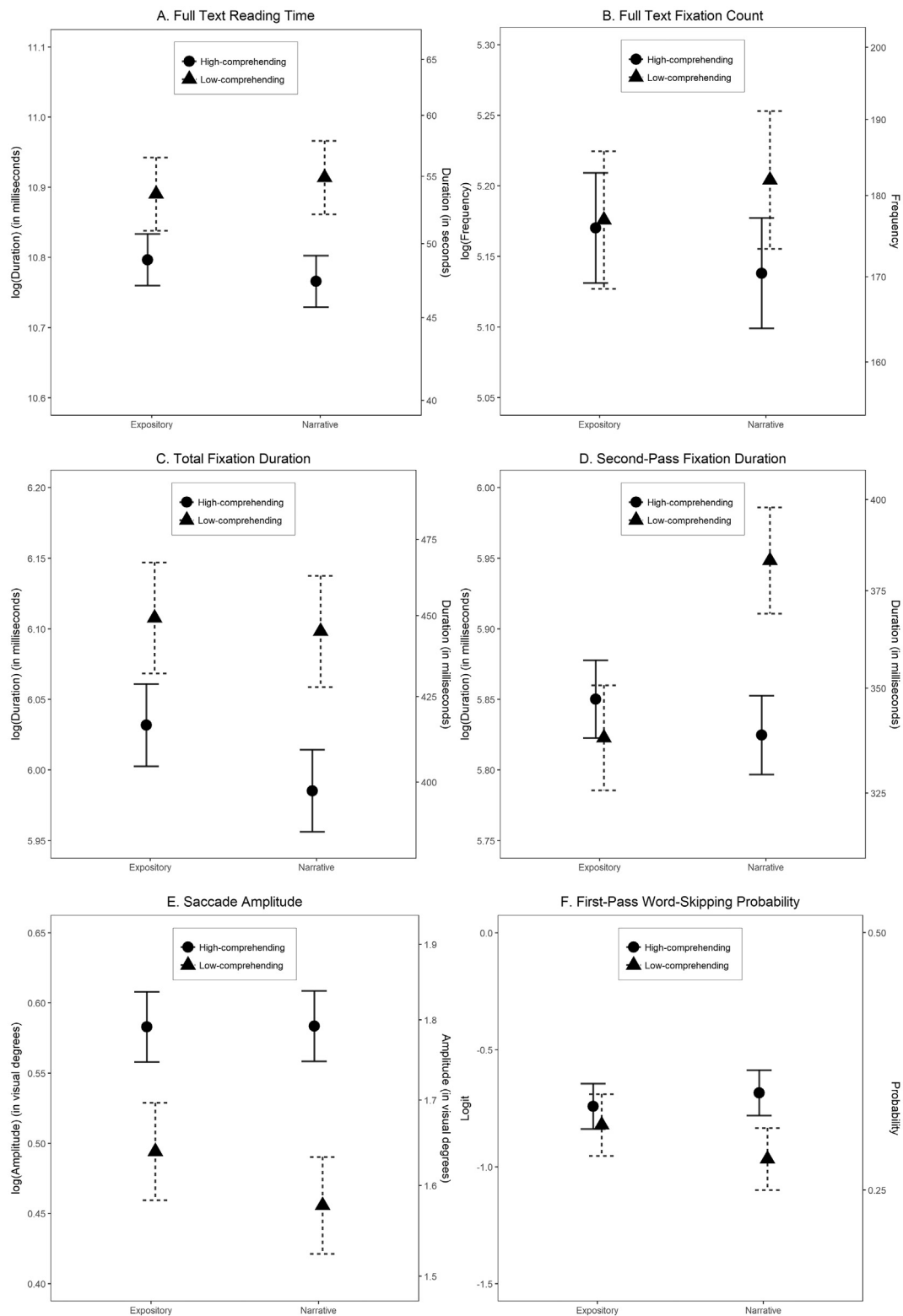


Fig. 1. Fixed effects estimates (and their SEs) of the eye-movements measures revealing a significant interaction between Text Genre (expository vs. narrative texts) and Reading Proficiency (high- vs. low-comprehending readers). Reading-time scales (for the log-transformed continuous measures) and logit scales (for the probability measures) are added on the y-axes for convenience.

texts in comparison to narrative texts ($\beta = 0.031$, $SE = 0.025$, $t = 1.23$), whereas low-comprehending readers displayed a tendency in the opposite direction ($\beta = -0.024$, $SE = 0.030$, $t = -0.78$).

3.3.1.2. Full text fixation count. The Wald chi-square tests revealed a significant interaction between Text Genre and Reading Proficiency ($\chi^2(1) = 5.35$, $p < 0.05$). The results for the simple effects reported below should be interpreted with caution, because only the interaction effect itself was reliable. That said, the pattern in Fig. 1B

suggests that low-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}185$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}32$) made more fixations than high-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}175$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}39$) while reading narrative texts ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.066$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.048$, $t\hat{a}\hat{c} = \hat{a}\hat{c}1.39$), but not while reading expository texts ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.0057$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.048$, $t\hat{a}\hat{c} = \hat{a}\hat{c}0.12$). As can be seen, this interaction occurred because high-comprehending readers made more fixations in expository texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}181$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}41$) than in narrative texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}175$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}39$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.032$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.043$, $t\hat{a}\hat{c} = \hat{a}\hat{c}0.74$), whereas low-comprehending readers showed the opposite pattern by making less fixations in expository texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}180$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}32$) than in narrative texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}185$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}32$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.028$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.046$, $t\hat{a}\hat{c} = \hat{a}\hat{c}0.61$).

3.3.1.3. First fixation duration. The Wald chi-square test revealed a main effect of Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}8.42$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.01$). Low-comprehending readers displayed longer first fixation durations than did high-comprehending readers ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.085$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.029$, $t\hat{a}\hat{c} = \hat{a}\hat{c}2.90$).

3.3.1.4. First gaze duration. The Wald chi-square tests revealed a main effect of Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}9.98$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.01$). Low-comprehending readers displayed longer first gaze durations than did high-comprehending readers ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.11$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.035$, $t\hat{a}\hat{c} = \hat{a}\hat{c}3.16$).

3.3.1.5. Right-bounded duration. The Wald chi-square tests revealed a main effect of Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}6.46$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.05$). Low-comprehending readers displayed longer right-bounded durations than did high-comprehending readers ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.11$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.041$, $t\hat{a}\hat{c} = \hat{a}\hat{c}2.54$).

3.3.1.6. Total fixation duration. The Wald chi-square tests revealed a main effect of Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}4.95$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.05$) and an interaction effect between Text Genre and Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}6.39$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.05$). Low-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}539$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}369$) displayed longer total fixation durations in narrative texts than did high-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}487$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}352$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.11$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.043$, $t\hat{a}\hat{c} = \hat{a}\hat{c}2.62$). This contrast between low- and high-comprehending readers was attenuated for expository texts ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.076$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.043$, $t\hat{a}\hat{c} = \hat{a}\hat{c}1.75$). Fig. 1C shows that this pattern emerged because low-comprehending readers did not display any notable difference in total fixation durations across expository ($M\hat{a}\hat{c} = \hat{a}\hat{c}545$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}391$) and narrative texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}539$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}369$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.0094$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.026$, $t\hat{a}\hat{c} = \hat{a}\hat{c}0.36$), whereas high-comprehending readers showed a marginal increase in total fixation durations while reading expository texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}517$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}398$), compared to while reading narrative texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}487$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}352$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.047$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.024$, $t\hat{a}\hat{c} = \hat{a}\hat{c}1.91$).

3.3.1.7. Second-pass fixation duration. The Wald chi-square tests revealed an interaction effect between Text Genre and Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}12.1$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.001$). Low-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}482.83$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}376.52$) displayed longer second-pass durations than did high-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}420.96$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}308.94$) while reading narrative texts ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.12$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.043$, $t\hat{a}\hat{c} = \hat{a}\hat{c}2.90$), whereas for expository texts this contrast between low-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}417.26$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}288.34$) and high-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}445.45$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}425.49$) was not present ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.027$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.042$, $t\hat{a}\hat{c} = \hat{a}\hat{c}0.65$). As illustrated in Fig. 1D, this interaction is primarily driven by decreasing second-pass durations for low-comprehending readers in expository texts ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.13$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.041$, $t\hat{a}\hat{c} = \hat{a}\hat{c}3.09$). In contrast, for high-comprehending readers there was no such difference in second-pass

durations between expository and narrative texts ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.025$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.032$, $t\hat{a}\hat{c} = \hat{a}\hat{c}0.80$).

3.3.1.8. Saccade amplitude. The Wald chi-square tests revealed a main effect of Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}8.16$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.01$) and an interaction effect between Text Genre and Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}4.17$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.05$). Low-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}2.14$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}2.02$) displayed smaller first-pass saccade amplitudes in narrative texts than did high-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}2.27$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}1.86$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.13$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.039$, $t\hat{a}\hat{c} = \hat{a}\hat{c}3.26$). For expository texts, this contrast between low-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}2.22$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}2.09$) and high-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}2.30$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}1.88$) was also present but in an attenuated manner ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.089$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.039$, $t\hat{a}\hat{c} = \hat{a}\hat{c}2.27$). As illustrated by Fig. 1E, this attenuation of the effect was due to a relatively stable saccade amplitude for high-comprehending readers across different text genres ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.00051$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.020$, $t\hat{a}\hat{c} = \hat{a}\hat{c}0.025$), as opposed to a (non-significant) increment of saccade amplitude for low-comprehending readers in expository texts ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.038$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.023$, $t\hat{a}\hat{c} = \hat{a}\hat{c}1.66$).

3.3.1.9. First-pass regression probability. The Wald chi-square tests revealed no reliable main or interaction effects.

3.3.1.10. First-pass word skipping probability. The Wald chi-square tests revealed a significant interaction between Text Genre and Reading Proficiency ($\hat{I}\hat{x}^2(1)\hat{a}\hat{c} = \hat{a}\hat{c}16.3$, $p\hat{a}\hat{c} < \hat{a}\hat{c}.001$). When reading narrative texts, low-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}0.28$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}0.45$) showed a trend towards skipping fewer words than did high-comprehending readers ($M\hat{a}\hat{c} = \hat{a}\hat{c}0.34$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}0.47$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.28$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.15$, $z\hat{a}\hat{c} = \hat{a}\hat{c}1.94$). This (marginally significant) simple effect of reading proficiency was less pronounced in expository texts ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.080$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.15$, $z\hat{a}\hat{c} = \hat{a}\hat{c}0.55$). As illustrated by Fig. 1F, the attenuation of the effect stemmed from the tendency of low-comprehending readers to skip more words while reading expository texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}0.32$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}0.47$) than while reading narrative texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}0.28$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}0.45$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.15$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.085$, $z\hat{a}\hat{c} = \hat{a}\hat{c}1.71$), as opposed to the tendency of high-comprehending readers in the opposite direction, i.e., less word-skipping in expository texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}0.33$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}0.47$) than in narrative texts ($M\hat{a}\hat{c} = \hat{a}\hat{c}0.34$, $SD\hat{a}\hat{c} = \hat{a}\hat{c}0.47$), ($b\hat{a}\hat{c} = \hat{a}\hat{c}0.057$, $SE\hat{a}\hat{c} = \hat{a}\hat{c}0.079$, $z\hat{a}\hat{c} = \hat{a}\hat{c}0.73$). Note, however, that the results for the simple effects should be interpreted with caution because only the interaction effect itself was reliable.

4. Discussion

In the present eye-tracking study, we explored the on-line text-processing approach of low- and high-comprehending second-grade readers as they read expository and narrative texts for comprehension. We hypothesized that, relative to narrative texts, beginning readers need to adjust their reading strategy for expository texts as a genre they are often unfamiliar with to optimize reading comprehension. Furthermore, we anticipated that high-comprehending readers adjust their reading strategy more easily (and therefore more noticeably) than low-comprehending readers do. Hence, we predicted that reading proficiency and text genre variables should interact and we expected a later eye-tracking measures associated with higher-level (e.g., top-down) integrative processing to be more sensitive to this interaction than an early eye-tracking measures associated with the processing stage of word identification and syntactic parsing.

Indeed, the results revealed evidence for reading proficiency and text genre interactions in several eye-movement metrics (i.e., full text reading time and fixation count, total fixation duration, second-pass

fixation duration, saccade amplitude, skipping probability). The general pattern was that narrative texts induced prolonged, more intensive processing for low-comprehending readers than for high-comprehending readers. This difference between low- and high-comprehending readers was attenuated in the case of expository texts. As we will discuss below, the directions of these interactions, however, are not always compatible with the idea that high-comprehending readers adapt their reading style to text genre whereas low-comprehending readers do not. We explore several scenarios that can account for the findings. We first discuss the findings with reference to strategic processing (i.e. along the lines of our hypotheses). We then introduce an alternative account that focuses on the interplay of working-memory capacity, prior knowledge, and the differential processing demands posed by narrative and expository texts.

4.1. Eye-movement profiles of low- and high-comprehending readers in narrative and expository texts

In the case of narrative texts, low-comprehending readers displayed longer full text reading times and average fixation durations per word than did high-comprehending readers; this was due, at least partly, to them making more fixations than high-comprehending readers did. For expository texts, the contrast between low- and high-comprehending readers diminished. The observed interaction for these relatively coarse measures of processing may reflect a strategic adjustment by high-comprehending readers (they slowed down in expository texts, relative to narrative texts) or by low-comprehending readers (they sped up in expository texts, relative to narrative texts), as the data are compatible with either account (see Fig. 1A-C). The analyses of the finer-grained (i.e., word-based) eye-movement measures provide a more detailed and somewhat mixed picture of why these interactions emerged.

The early eye-movement measures revealed that low-comprehending readers displayed longer first fixation and first gaze durations than high-comprehending readers did. As expected, this seems to indicate that low-comprehending readers experienced more problems with (bottom-up) processes of word identification and early (syntactic) integration (e.g., Boland, 2004). Somewhat surprisingly, although the expository texts were more difficult than the narrative texts (3rd vs. 2nd grade levels, respectively) we observed no significant main effects of genre in the early measures but note that the first gaze durations showed a numerical difference in the expected direction. Most importantly, we observed no interaction effects between reading proficiency and text genre. This seems to indicate that strategic adjustments of readers' processing style do not affect the early measures extracted from the eye-movement data.

Later eye-movement measures that likely reflect more strategic processing painted a somewhat murky picture. A valid conclusion from the absence of significant genre differences for high-comprehending readers would be that they employed a relatively stable processing approach across narrative and expository texts. However, this conclusion does not fit well with the overall pattern for high-comprehending readers as nearly all (untransformed) measures show a numerical difference in the expected direction: high-comprehending readers seem to slow down (i.e. adjust) their reading pace when confronted with an expository text (see Table 5). The interpretation of the results for the low-comprehending readers is not clear-cut either. Again, if we base our conclusion strictly on statistically significant results, low-comprehending readers demonstrated a relatively stable processing approach i.e., only second-pass durations differed significantly across narrative and expository texts for these readers. That being said, there are also signs that low-comprehending readers adapted their reading behavior to text genre, yet in an unexpected way. First, the analyses for second-pass durations showed that if low-comprehending readers reread a word, they were inclined to spend less time rereading the words of expository texts than the words of narrative texts. Second, similar patterns emerged for measures of saccade amplitude and word-skipping

probability. Again, low-comprehending readers showed an inclination towards a less intensive processing approach for expository texts (i.e., longer saccades, more word skipping). We should emphasize, however, that the interpretation for these latter two measures is complicated by the fact that follow-up analyses of the simple effects returned marginally significant or non-significant results. Therefore, although visual inspection of the significant interactions reveals a consistent picture, we can only be tentative in concluding that low-comprehending readers adjusted their reading strategy whereas high-comprehending readers did not.

So far, we have assumed that the interaction effects have emerged as a result of differential reading strategies by low- and high-comprehending pupils (i.e., we speculated that for expository texts high-comprehending readers have a tendency towards a relatively intensive processing approach, whereas low-comprehending readers have a tendency towards a more cursory processing approach). Hence, our take on the results focuses on the (top-down) influence of readers and not so much on the characteristics of a text. It is also possible that the observed interaction effects are primarily driven by the text, specifically by what a reader knows about the content and structure of the text prior to reading this text. In general, expository texts are used for learning and most readers approach the task of reading these texts with low levels of topic-relevant knowledge. The reading process of narrative texts, in contrast, is supported by sufficient knowledge about the content and, as a result, within- and between-sentence inferences may be generated relatively automatically (Best, Rowe, Ozuru, & McNamara, 2005). So, one way to approach the reading advantage of high-comprehending readers compared to low-comprehending readers for narrative texts is to assume that high-comprehending readers have more prior knowledge about the content of narrative texts and that the inferences generated in these texts are carried out more in an adult-like manner, requiring only a limited amount of cognitive resources. These advantages would not hold for expository texts because the content will be new to both low- and high comprehending readers and, hence, inferences will be resource-consuming to both types of readers. As pointed out by an anonymous reviewer, however, another possibility is that the processing of narrative texts relies less on prior knowledge than the processing of expository texts because narrative texts can unfold in many different, unpredictable ways. Consequently, narrative texts could make a greater demand on working-memory resources because "new", event-like information needs to be kept active during reading (cf., Williams et al., 2004). As indexed by the eye-tracking data, this would primarily affect the group of low-comprehending, low verbal working-memory capacity (see Table 3) readers, while they process narrative texts. Overall, for both narrative and expository texts the conjecture can be made that readers will be unfamiliar with their content, which could explain why high-comprehending readers display a processing advantage (i.e., as indicated by the eye-movement metrics) for narrative texts, but less so for expository texts.

4.2. Coping strategies of low-comprehending readers

A general conclusion that we would like to draw from this study is that, relative to narrative texts, low-comprehending readers reacted somewhat differently to expository texts than high-comprehending readers did. The results do not allow for a definitive conclusion about the source(s) of this interaction, however. It may be that low-comprehending readers adopted a more cursory processing approach for expository texts or, alternatively, that high-comprehending readers displayed a more intensive processing approach for expository texts or these tendencies may both underlie the observed interactions. The results do point out that low-comprehending readers adopt a suboptimal processing approach for expository texts because expository texts tend to demand more cognitive effort of the reader than narrative texts (Williams et al., 2004). In this light, an effortful or intensive reading style seems more appropriate.

If low-comprehending readers have a relatively stable processing approach across different text genres this would suggest that, although they may have acquired an effective reading routine for narrative texts, these eight-year-old readers still have to develop more optimal (intensive) reading strategies for expository texts and learn how to employ them flexibly in different situations. If, however, low-comprehending readers adopt a more cursory processing approach for expository texts, we should provide an explanation for why they opt for such a processing approach.

One possible explanation is that low-comprehending readers had lower standards of coherence for expository texts than for narrative texts. For more experienced readers, expository texts tend to evoke strict standards of coherence because readers associate them with learning goals, whereas narrative texts tend to evoke more shallow standards of coherence because readers associate them with reading for pleasure (Van den Broek et al., 2011). For younger (low-comprehending) children this contrast may not be present, or may even be reversed, because they tend to read many narrative texts in an educational setting. Moreover, low-comprehending pupils may not be capable of applying strict standards of coherence, because they lack the required monitoring skills (Oakhill et al., 2005) or may simply not strive for full comprehension of the text because they view reading mainly as a decoding activity (Cain, 1999; Yuill & Oakhill, 1991). Our observation that low-comprehending readers decreased their refixation duration (and more speculatively, made longer saccades and tended to skip more words) when reading expository texts supports the idea that they lowered their standards of coherence for these texts. However, low-comprehending readers scored relatively high on knowledge-based questions for expository texts. This suggests that they still attempted to relate the text to their background knowledge, which is not what one would expect from a reader who had lowered his or her standards altogether.

A second possibility is that for expository texts low-comprehending readers engaged in a so-called "exploring processing approach" (cf., Olson et al., 1985). This would mean that low-comprehending readers did not necessarily have low standards of coherence for expository texts, but they appealed to their background knowledge to compensate for comprehension difficulties (Coté et al., 1998). However, they did not make the regressive eye-movements associated with an exploring text-processing approach (Olson et al., 1985; cf. Koornneef & Mulders, 2016; Rayner, Castelano, & Yang, 2009). Regressive eye-movements can be an indication of effective comprehension-monitoring skills (Ehrlich et al., 1999; Rayner et al., 2016; Schotter et al., 2014; Van der Schoot et al., 2012), but as pointed out by Pearson et al. (1992), low-comprehending readers often lack adequate repair strategies (e.g., looking back to earlier sections of a text) to adapt their reading approach in an effective manner when comprehension fails.

A third possibility is that the cursory processing approach was the only viable strategy low-comprehending readers possessed for coping with expository texts, because they simply lacked the "tools" to engage in more effortful processing (Denton et al., 2015). This may be caused by various reader characteristics, or a combination of characteristics, such as lack of sufficient word and/or world knowledge (Coté et al., 1998; Denton et al., 2015), insufficient inference and integration skills (Cain & Oakhill, 1999), and limited working-memory capacity (Linderholm & Van den Broek, 2002). In this context, it should be reiterated that the low-comprehending readers in the present study obtained lower scores than high-comprehending readers on the tests for receptive word knowledge, inference making, and verbal working-memory capacity.

Finally, a fourth possibility is that the results in the current study emerged due to issues that are unrelated to the strategic (cursory versus intensive) processing styles of readers. As outlined in Section 4.1, "optimal" reading behavior is related not only to the strategic enactment of processing but also to what a reader already knows about the topic of a text. This especially holds for expository texts (i.e., â

facts are facts"™), which may explain why low- and high-comprehending readers behave similarly while reading expository texts (i.e., because their prior knowledge of the topics does not differ much) but not while reading narrative texts (i.e., prior knowledge is less important in narrative texts because these texts tend to describe new events).

These alternative interpretations of the current findings are worth comparing directly in future research. Such research could also vary methodological aspects of the current study that both allow and limit the conclusions. For example, the topics and questions of the texts were not matched across genre (hence, main effects of text genre are difficult to interpret). Furthermore, we presented the texts in the same fixed order to all participants. This has the advantage that the impact of individual differences between readers on their eye-movement routines can be examined more adequately. However, this also meant that the narrative texts were read later in the experiment (texts 2 and 4) than the expository texts (texts 1 and 3). As a result, the main and interaction effects in the eye-movement metrics and comprehension questions may in part reflect the potential influence of fatigue and/or familiarity with the experimental procedure. Finally, due to strict selection criteria for low-comprehending readers (i.e., average or above-average scores on standardized tests for word-reading ability and non-verbal intelligence) the sample size of this group was relatively small. These limitations of the study should be kept in mind when considering potential implications for education in the following sections.

4.3. Educational implications

Based on our interpretation of the data, we suggest that low-comprehending readers in the second grade of primary school employ suboptimal reading strategies for expository texts. On the one hand, if low-comprehending readers adopt a relatively stable processing approach across text genre, they are likely to benefit most from interventions aimed at clarifying the differences between narrative and expository texts and stimulating a more intensive processing style for the latter. On the other hand, if low-comprehending readers adopt a more cursory processing approach for expository texts, the educational implications are less straightforward. We suggested that a cursory processing strategy may stem from lower standards of coherence, an exploring processing approach, or a coping strategy. Depending on which of these explanations best describes the processing strategy of low-comprehending readers, the educational implications will differ.

If low-comprehending readers indeed have low standards of coherence for expository texts, this is not desirable in educational settings because expository texts are widely used to transfer knowledge. In that case it could be useful to make children aware of their coherence standards, e.g., by discussing the task and the reading goals before reading (Cain & Oakhill, 2007). Furthermore, when the attention of readers is guided to certain information in a text that they have processed insufficiently, this can lead to deeper text processing, stricter standards of coherence, and better comprehension (Sanford, Sanford, Molle, & Emott, 2006).

If the correct interpretation is that low-comprehending readers adopted an exploring processing style for expository texts "comparable to Olson et al.'s (1985) explorers, but without the characteristic regressions" this will present different challenges and may require a different approach on the part of educational practitioners. A positive aspect of this reading strategy is that the readers are attempting to connect the information in the text to their background knowledge. There are risks in this, however: It may result in the reader learning less new information from texts than he or she would derive with a more text-based reading manner (Coté et al., 1998), especially if the background knowledge is insufficient or incorrect. In the latter case, misconceptions about a topic will easily persist without a proper intervention (Coté et al., 1998). For this reason, when readers have little or erroneous background knowledge, a more effortful processing style seems more appropriate.

If the third explanation is correct, and low-comprehending readers' reading profile for expository texts reflects an optimal coping strategy, it remains to be seen whether they would benefit from a more effortful reading approach for such texts. Indeed, it is plausible that the mental text representations of low-comprehending readers may become more coherent if they make more and stronger text-connecting inferences, for instance. However, they need to be capable of doing so. In this case, rather than trying to adjust the global reading style of low-comprehending readers from a cursory to a more effortful style, it might be more useful to address the underlying cognitive and language deficiencies.

4.4. Conclusion

Investigation of individual differences in text processing across different text genres can provide important insights into comprehension differences; these insights may eventually help low-comprehending readers to overcome their reading difficulties and high-comprehending readers to further improve their reading skills. Our study showed that narrative texts induce prolonged, more intensive processing for low-comprehending readers than for high-comprehending readers. This difference between low- and high-comprehending readers was attenuated during reading of expository texts. Accordingly, we speculated that either low-comprehending readers adopted a more

cursory processing approach for expository texts or that high-comprehending readers displayed a more intensive processing approach for expository texts or both. In our opinion the results show that, in any case, low-comprehending readers adopt a suboptimal processing approach for expository texts, by either not adjusting their reading strategy or by adopting a more cursory strategy. This raises important questions about how we can best help young (low-comprehending) readers to become effective, flexible and strategic readers.

Funding

This work was supported by the Netherlands Organization for Scientific Research (NWO, grant BGRL-11-09).

Declaration of Competing Interest

None.

Acknowledgements

The authors wish to thank Prof. Dr. Huestegge and three anonymous reviewers for their valuable feedback and input on this paper, and Esther Brussaard, Willeke Buys, Simone Hemerik, Ghislaine Le Large and Anne-Karin Lok for their assistance in collecting the data.

Appendix A. Original version of the experimental texts and comprehension questions and their translations into English

#1 Expository text "Skeleton"

<p>In je lijf zitten botten. Alle botten samen zijn je skelet. Veel botten kun je voelen. Bijvoorbeeld die in je vingers en je tenen. Botten zijn gemaakt van kalk en lijmfstof. Het ene maakt je botten sterk. En het andere maakt ze buigbaar. Kleine kinderen zijn heel lenig. Er zit veel lijmfstof in hun botten. Hun botten breken niet snel. Bij oude mensen is het anders. Zij hebben weinig lijmfstof in hun botten. Er zit meer kalk in hun botten. Hun botten breken sneller. Je skelet geeft vorm aan je lijf. Het bestaat uit meer dan tweehonderd botten. Weet je waar de meeste botten zitten? In je handen en in je voeten. En het grootste bot in je lijf zit in je bovenbeen.</p> <p><i>Literal questions:</i> (#1) Waar zijn botten van gemaakt? (#6) Waar zit je grootste bot?</p> <p><i>Text-based questions requiring a text-connecting inference (between sentences):</i> (#2) Waarom breken botten van kleine kinderen niet snel? (#3) Waarom breken botten van oude mensen sneller? (#5) Waar zitten de meeste botten?</p> <p><i>Knowledge-based question requiring a gap-filling inference:</i> (#4) Hoe denk dat je lijf zou zijn zonder botten, zonder skelet?</p>	<p>Your body contains bones. The bones together form your skeleton. Many bones you can feel. For example, those in your fingers and your toes. Bones are made of calcium and collagen. One makes your bones strong. And the other makes them bendable. Small children are very flexible. There is a lot of collagen in their bones. Their bones do not break easily. It is different for old people. They have little collagen in their bones. There is more calcium in their bones. Their bones break more easily. Your skeleton gives shape to your body. It consists of more than two hundred bones. Do you know where most bones are? In your hands and in your feet. And the biggest bone in your body is in your thigh.</p> <p>What are bones made of? Where in your body is your biggest bone?</p> <p>Why do small children's bones not break easily? Why do old people's bones break easily? Where are most bones in your body?</p> <p>What do you think: What would your body be like without skeleton?</p>
--	--

#2 Narrative text "iPad"

<p>Papa is eten aan het koken. Mama zit te werken achter de computer. Simon en Tom spelen met de iPad van hun grote zus. Die ligt een boek te lezen op de bank. Maar dan wordt opeens het scherm zwart. De iPad doet helemaal niets meer. Trillend lopen ze naar hun zus. Ze wijzen naar de iPad. Simon zegt: Hij ging opeens kapot.</p>	<p>Dad is cooking dinner. Mama is working at the computer. Simon and Tom are playing with their big sister's iPad. She is reading a book on the couch. But then suddenly the screen turns black. The iPad does nothing at all anymore. Trembling, they walk over to their sister. They point to the iPad. Simon says: It broke suddenly.</p>
--	--

Sanne kijkt haar broertjes aan en begint te lachen.
 Het valt wel mee hoor!, zegt ze.
 Zullen we hem weer maken?
 Sanne pakt een oplader uit de kast.
 En ze zegt: Na het eten doet hij het weer.
 De batterij is alleen maar leeg.
 Hij moet even in de oplader.
 Tom en Simon zijn opgelucht.
 Papa roept dat ze allemaal moeten komen.
 Zo, nu eerst lekker eten.

Literal questions:

(#4) Wat pakt Sanne uit de kast?

(#5) Wie roept dat ze moeten komen?

Text-based questions requiring a text-connecting inference (between sentences):

(#1) Waarom denken Simon en Tom dat de iPad kapot is?

(#2) Hoe heet de zus van Simon en Tom?

(#3) Waarom begint Sanne te lachen?

Knowledge-based question requiring a gap-filling inference:

(#6) Wanneer op de dag speelt dit verhaal zich af, denk je?

Sanne looks at her brothers and starts to laugh.
 It is not that bad!, she says.
 Shall we fix it?
 Sanne takes a charger out of the closet.
 And she says: After dinner, it will work again.
 The battery is just empty.
 It needs to be charged.
 Tom and Simon are relieved.
 Papa calls for them to come.
 Let's eat dinner first.

What does Sanne take out of the closet?

Who calls them telling them to come?

Why do Simon and Tom think the iPad is broken?

What is the name of the sister of Simon and Tom?

Why does Sanne start to laugh?

What do you think: What time of day does this story take place?

#3 Expository text "Lions"

Leeuwen leven in een groep.

Zo'n groep leeuwen heet een troep.

Het zijn ongeveer vijftien leeuwen bij elkaar.

Het is handig om in een troep te leven.

De leeuwen jagen samen.

Zo hebben ze meer kans om een prooi te vangen.

En ze zorgen samen voor de jongen.

Die hebben zo meer kans om te overleven.

Een mannetjesleeuw heeft manen.

Dat zijn lange haren op zijn kop en nek.

Daarmee valt hij erg op.

Een mannetje kan heel hard brullen.

Zo jaagt hij andere leeuwen weg.

Bij leeuwen moeten de vrouwtjes jagen.

Zij zorgen voor het eten.

Het mannetje jaagt niet mee.

Mannetjes vallen veel te veel op.

Van jagen worden leeuwen moe.

Ze slapen veel, soms wel twintig uur per dag.

Literal questions:

(#1) Wat is een troep?

(#5) Wie moeten jagen bij leeuwen?

Text-based questions requiring a text-connecting inference (between sentences):

(#2) Waarom is het handig om in een troep te jagen?

(#4) Hoe komt het dat mannetjes te veel opvallen?

(#6) Waarom slapen leeuwen twintig uur per dag?

Knowledge-based question requiring a gap-filling inference:

(#3) Waarom denk je dat jongen in een groep meer kans hebben om te overleven?

Lions live in a group.

A group of lions is called a troop.

It is about fifteen lions together.

It is useful to live in a troop.

The lions hunt together.

That way, they are more likely to catch their prey.

And they take care of the young together.

So the young more likely to survive.

A male lion has a mane.

Those are long the hairs on his head and neck.

He is therefore very striking.

A male can roar very loudly.

That is how he chases off other lions.

For lions, the females have to do the hunting.

They provide the food.

The male does not hunt.

Males stand out too much.

Hunting makes lions tired.

They sleep a lot, sometimes as much as twenty hours a day.

What is a troop?

For lions, who does the hunting?

Why is it useful to hunt in a troop?

Why do male lions stand out?

Why do lions sleep twenty hours a day?

Why, do you think, the chances of survival for the young are better in a troop?

#4 Narrative text "Hide-and-Seek"

Alle kinderen zijn buiten want het is pauze.

Stef doet verstopperij met zijn vrienden.

Luuk heeft zich heel goed verstopt.

Hij is in een hoge boom geklommen.

Stef moet heel lang zoeken.

Als eerste vindt hij Julia.

Die had zich in het fietsenhok verstopt.

Zij is af en moet aan de kant zitten.

Stef zoekt verder en verder.

Bijna iedereen heeft hij gebuut.

Alleen Luuk kan hij niet vinden.

Julia ziet Luuk wel.

Ze kijkt omhoog naar Luuk.

Stef ziet Julia omhoog kijken.

En dan ziet Stef hem ook.

Heel blij roept hij meteen: Buut Luuk!

Luuk vindt het niet eerlijk.

Hij is een beetje boos op Julia.

Dan gaat de bel en moeten ze naar binnen.

The children are outside because it is recess.

Stef is playing hide-and-peek with his friends.

Luuk is well hidden.

He has climbed into a tall tree.

Stef has to search for a long time.

First, he finds Julia.

She had hidden in the bicycle shed.

She has been found and has to sit on the side.

Stef searches on and on.

He finds almost everyone.

He just cannot find Luuk.

But Julia sees Luuk.

She looks up at Luuk.

Stef sees Julia looking up.

And then Stef also sees him.

Happily he immediately yells: I see you, Luuk!

Luuk does not think it is fair.

He is a little angry with Julia.

Then the bell rings and they have to go inside.

Literal questions:

- (#1) Welk spel doen Stef en zijn vrienden?
 (#3) Wie vindt Stef als eerste?

What game are Stef and his friends playing?
 Who does Stef find first?

Text-based questions requiring a text-connecting inference (between sentences):

- (#2) Waar zit Luuk verstopt?
 (#4) Waarom ziet Stef Luuk?
 (#5) Wat vindt Luuk niet eerlijk?

Where is Luuk hiding?
 Why does Stef see Luuk?
 Why does Luuk think it's not fair?

Knowledge-based question requiring a gap-filling inference:

- (#6) Waar denk je dat Luuk en zijn vrienden zijn?

Where do you think Luuk and his friends are?

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious mixed models*. arXiv preprint arXiv:1506.04967.
- van Berkel, S., Engelen, R., van Groen, M., Hilde, M., Wouda, J., & Van der Zanden, J. (2013). *Wetenschappelijke verantwoording Begrijpend luisteren groep*. Cito B.V.: Arnhem:3.
- Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, 29(2), 137–164. <https://doi.org/10.1080/02702710801963951>.
- Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders*, 25(1), 65–83. <https://doi.org/10.1097/00011363-200501000-00007>.
- Blythe, H. I., HÅikiÅ, T., Bertam, R., Liversedge, S. P., & HyÅnÅ, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research*, 51, 84–92. <https://doi.org/10.1016/j.visres.2010.10.003>.
- Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular co-ordination of eye movements during reading in children and adults. *Vision Research*, 46, 3898–3908. <https://doi.org/10.1016/j.visres.2006.06.006>.
- Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). The uptake of visual information during fixations in reading in children and adults. *Vision Research*, 49, 1583–1591. <https://doi.org/10.1016/j.visres.2009.03.015>.
- Boland, J. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras, & C. Clifton Jr. (Eds.). *The on-line study of sentence comprehension* (pp. 51–76). New York: Psychology Press.
- Boston, M. F., Hale, J. T., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, 2(1).
- Cain, K. (1999). Ways of reading: How knowledge and use of strategies are related to reading comprehension. *British Journal of Developmental Psychology*, 17(2), 293–309. <https://doi.org/10.1348/026151099165285>.
- Cain, K., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal*, 11, 489–503. <https://doi.org/10.1023/A:1008084120205>.
- Cain, K., & Oakhill, J. (2007). Reading comprehension difficulties: Correlates, causes, and consequences. In K. Cain, & J. Oakhill (Eds.). *Children's comprehension problems in oral and written language: A cognitive perspective* (pp. 41–75). New York, NY: Guilford.
- Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference making ability and their relation to knowledge. *Memory and Cognition*, 29, 850–859. <https://doi.org/10.3758/BF03196414>.
- Cito (2006). *LOVS Begrijpend lezen groep 3 [Reading Comprehension Grade 1]*. Arnhem: Cito.
- Cito (2009). *Drie-Minuten-Test [Three-Minutes-Test]*. Arnhem: Cito.
- Cito (2011). *LOVS Begrijpend Luisteren groep 3 [Listening Comprehension Grade 1]*. Arnhem: Cito.
- Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.). *Eye movements: A window on mind and brain* (pp. 341–372). Oxford, UK: Elsevier.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769–786. <https://doi.org/10.3758/BF03196772>.
- CotÅ, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, 25(1), 1–53. <https://doi.org/10.1080/01638539809545019>.
- Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., & Carter, S. (2015). Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly*, 50(4), 393–416. <https://doi.org/10.1002/rq.105>.
- 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. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup, & S. Jay Samuels (Eds.). *What research has to say about reading instruction* (pp. 205–242). (3rd ed.). Newark, DE: International Reading Association, Inc. <https://doi.org/10.1598/0872071774.10>.
- Ehrlich, M. F., Remond, M., & Tardieu, H. (1999). Processing of anaphoric devices in young skilled and less skilled comprehenders: Differences in metacognitive monitoring. *Reading and Writing*, 11(1), 29–63. <https://doi.org/10.1023/A:1007996502372>.
- Evers, G. Programma voor berekening Cito LeesIndex voor het Basisonderwijs. P-CLIB versie 3.0. Arnhem: Cito, 1994–2008.
- Feenstra, H., Kamphuis, F., Kleintjes, F., & Krom, R. (2010). *Wetenschappelijke verantwoording Begrijpend lezen voor groep 3 tot en met 6*. Arnhem: Cito B.V. Retrieved September 23, 2018, from <http://docplayer.nl/370648-Wetenschappelijke-verantwoording-begrijpend-lezen-voor-groep-3-tot-en-met-6.html>.
- Helder, A., Van den Broek, P., Van Leijenhorst, L., & Bekker, K. (2013). Sources of comprehension problems during reading. In B. Miller, L. Cutting, & P. McCardle (Eds.). *Unraveling reading comprehension: Behavioral, neurobiological, and genetic components* (pp. 43–53). Baltimore, MD: Paul H. Brookes Publishing.
- Huestegge, L. (2010). Effects of vowel length on gaze durations in silent and oral reading. *Journal of Eye Movement Research*, 3(5), <https://doi.org/10.16910/jemr.3.5.5>.
- Huestegge, L., & Bocianski, D. (2010). Effects of syntactic context on eye movements during reading. *Advances in Cognitive Psychology*, 6, 79. <https://doi.org/10.2478/v10053-008-0078-0>.
- Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, 49, 2948–2959. <https://doi.org/10.1016/j.visres.2009.09.012>.
- Joseph, H. S., Bremner, G., Liversedge, S. P., & Nation, K. (2015). Working memory, reading ability and the effects of distance and typicality on anaphor resolution in children. *Journal of Cognitive Psychology*, 27(5), 622–639. <https://doi.org/10.1080/20445911.2015.1005095>.
- Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, 49, 2078–2086. <https://doi.org/10.1016/j.visres.2009.05.015>.
- Juhász, B. J., & Pollatsek, A. (2011). Lexical influences on eye movements on reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.). *The Oxford handbook of eye movements* (pp. 643–662). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199539789.013.0048>.
- Kaakinen, J. K., HyÅnÅ, J., & Keenan, J. M. (2003). How prior knowledge, WMC, and relevance of information affect eye fixations in expository text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(3), 447. <https://doi.org/10.1037/0278-7393.29.3.447>.
- Kaakinen, J. K., Lehtola, A., & Paatlammi, S. (2015). The influence of a reading task on children's eye movements during reading. *Journal of Cognitive Psychology*, 27(5), 640–656. <https://doi.org/10.1080/20445911.2015.1005623>.
- Koornneef, A. W., & Mulders, I. C. M. C. (2016). Can we read the eye-movement patterns of readers? Unraveling the relationship between reading profiles and processing strategies. *Journal of Psycholinguistic Research*, 45(1), 1–18. <https://doi.org/10.1007/s10936-016-9418-2>.
- Kraal, A., Koornneef, A. W., Saab, N., Broek, P. W., & Den, Van (2018). Processing of expository and narrative texts by low- and high-comprehending children. *Reading and Writing*, 31(9), 2017–2040. <https://doi.org/10.1007/s11145-017-9789-2>.
- Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). *Wetenschappelijke verantwoording DMT en AVI*. Arnhem: Cito. Retrieved September 23, 2018, from <https://docplayer.nl/9490037-Wetenschappelijke-verantwoording-dmt-en-avi.html>.
- Kuhlemeier, H., Jolink, A., KrÅmer, I., Hemker, B., Jongen, I., van Berkel, S., & Bechger, T. (2014). *Balans van de leesvaardigheid in het basis- en speciaal basisonderwijs 2*. *Uitkomsten van de peilingen in 2011 en 2012 in groep 8, groep 5 en de eindgroep van het SBO (PPON-reeks nummer 54)*. Arnhem: Cito.
- Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, 65(1), 42–73. <https://doi.org/10.1016/j.jml.2011.03.002>.
- de Leeuw, L., Segers, E., & Verhoeven, L. (2016). The effect of student-related and text-related characteristics on student's reading behaviour and text comprehension: An eye movement study. *Scientific Studies of Reading*, 20(3), 248–263. <https://doi.org/10.1080/10888438.2016.1146285>.
- Leeuw, L., Segers, E., & Verhoeven, L. (2016). Role of text and student characteristics in real-time reading processes across the primary grades. *Journal of Research in Reading*, 39(4), 389–408. <https://doi.org/10.1111/1467-9817.12054>.
- Linderholm, T., & Van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology*, 94, 778–784. <https://doi.org/10.1037/0022-0663.94.4.778>.
- Lorch, R. F., Jr. (2017). What is so difficult about expository text? In J. A. LeÅn, & I. Escudero (Eds.). *Reading comprehension in educational settings* (pp. 145–167). John Benjamins: Amsterdam.
- McDaniel, M. A., & Einstein, G. O. (1989). Material-appropriate processing: A

- contextualist approach to reading and studying strategies. *Educational Psychology Review*, 1, 113–145. <https://doi.org/10.1007/BF01326639>.
- Meyer, B. J., & 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., & 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. (1975). *The organization of prose and its effects on memory*. Amsterdam: North Holland.
- Meyer, B. J. F. (1985). Prose analysis: Purposes, procedures and problems. In B. K. Britton, & J. Black (Eds.). *Analyzing and understanding expository text* (pp. 11–64, 269–304). Hillsdale, NJ: Erlbaum.
- Nation, K. (2005). Children's reading comprehension difficulties. In M. J. Snowling, & C. Hulme (Eds.). *The science of reading* (pp. 248–265). Oxford: Blackwell.
- National Center for Education Statistics (2011). The nation's report card: Reading 2011. Retrieved May 1, 2014, from <https://nces.ed.gov/nationsreportcard/pdf/main2011/2012457.asp>.
- Oakhill, J., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, 16(2), 91–121. <https://doi.org/10.1080/10888438.2010.529219>.
- Oakhill, J., Cain, K., & Elbro, C. (2014). *Understanding and teaching reading comprehension: A handbook*. London: Routledge.
- Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing*, 18(7), 657–686. <https://doi.org/10.1007/s11145-005-3355-z>.
- Oakhill, J. V., & Cain, K. (2007). Issues of causality in children's reading comprehension. In D. McNamara (Ed.). *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 47–72). New York: Erlbaum.
- Oakland, T., & Lane, H. B. (2004). Language, reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing*, 4(3), 239–252. https://doi.org/10.1207/s15327574ijt0403_3.
- Olson, R. K., Kliegl, R., Davidson, B. J., & Foltz, G. (1985). Individual differences and developmental differences in reading disability. In G. MacKinnon, & T. G. Waller (Eds.). *Reading research: Advances in theory and practice* (pp. 1–64). New York: Academic Press.
- Pearson, P. D., Roehler, L., Dole, J., & Duffy, G. (1992). Developing expertise in reading comprehension. In S. J. Samuels, & A. E. Farstrup (Eds.). *What research has to say about reading instruction* (pp. 145–199). (2nd ed.). Newark, DE: International Reading Association.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling, & C. Hulme (Eds.). *The science of reading: A handbook* (pp. 227–247). Oxford, UK: Blackwell. <https://doi.org/10.1002/9780470757642.ch13>.
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Pressley, M., & Wharton-McDonald, R. (1997). Skilled comprehension and its development through instruction. *School Psychology Review*, 26(3), 448–467.
- Rapp, D. N., Van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in poor readers: A perspective for research and intervention. *Scientific Studies of Reading*, 11(4), 289–312. <https://doi.org/10.1080/10888430701530417>.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales. Section I: General overview*. San Antonio, TX: Harcourt Assessment.
- Ray, M. N., & Meyer, B. J. (2011). Individual differences in children's knowledge of expository text structures: A review of literature. *International Electronic Journal of Elementary Education*, 4(1), 67–82.
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, 41(2), 211–236. [https://doi.org/10.1016/0022-0965\(86\)90037-8](https://doi.org/10.1016/0022-0965(86)90037-8).
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>.
- Rayner, K., Castelano, M. S., & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging*, 24(3), 755. <https://doi.org/10.1037/a0014300>.
- Rayner, K., Schotter, E. R., Masson, M., Potter, M. C., & Treiman, R. (2016). So much to read, so little time: How do we read, and can speed reading help? *Psychological Science in the Public Interest*, 17(1), 4–34. <https://doi.org/10.1177/1529100615623267>.
- Sanford, A. J. S., Sanford, A. J., Molle, J., & Emott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes*, 42(2), 109–130. https://doi.org/10.1207/s15326950dp4202_2.
- Schlichting, L. (2005). *Peabody picture vocabulary test-III-NL*. Amsterdam: Harcourt Test Publisher.
- Schotter, E. R., Tran, R., & Rayner, K. (2014). Don't believe what you read (only once) comprehension is supported by regressions during reading. *Psychological Science*, 25(6), 1218–1226.
- Schroeder, S. (2011). What readers have and do: Effects of students' verbal ability and reading time components on comprehension with and without text availability. *Journal of Educational Psychology*, 103(4), 877. <https://doi.org/10.1037/a0023731>.
- Swanson, H. L., Cochran, K. F., & Ewers, C. A. (1989). Working memory in skilled and less skilled readers. *Journal of Abnormal Child Psychology*, 17, 145–156. <https://doi.org/10.1007/BF00913790>.
- Traxler, M. J., Bybee, M. D., & Pickering, M. J. (1997). Influence of connectives on language comprehension: Eye tracking evidence for incremental interpretation. *The Quarterly Journal of Experimental Psychology: Section A*, 50(3), 481–497. <https://doi.org/10.1080/027249897391982>.
- Van den Broek, P., Bohn-Gettler, C., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. Mc Crudden, J. Magliano, & G. Schraw (Eds.). *Relevance instructions and goal-focusing in text learning* (pp. 123–140). Greenwich, CT: Information Age Publishing.
- Van den Broek, P., & Espin, C. A. (2012). Connecting cognitive theory and assessment: Measuring individual differences in reading comprehension. *School Psychology Review*, 41, 315–325.
- Van den Broek, P., & Kendeou, P. (2017). In K. Cain, & D. Compton (Eds.). *The development of language comprehension skills: Change and continuity in the ability to build coherence* Amsterdam: John Benjamins Theories Of Reading Development, a 20 year celebration of Scientific Studies of Reading.
- Van den Broek, P., White, M. J., Kendeou, P., & Carlson, S. (2009). Reading between the lines: Developmental and individual differences in cognitive processes in reading comprehension. In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.). *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 107–123). New York: Guilford.
- Van der Schoot, M., Reijntjes, A., & Van Lieshout, E. C. M. D. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing*, 25(7), 1665–1690. <https://doi.org/10.1007/s11145-011-9337-4>.
- Vorstius, C., Radach, R., & Lonigan, C. J. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition*, 22(3&4), 458–485. <https://doi.org/10.1080/13506285.2014.881445>.
- Williams, J. P., Hall, K. M., & Lauer, K. D. (2004). Teaching expository text structure to young at-risk learners: Building the basics of comprehension instruction. In S. Vaughn, & J. P. Williams (Vol. Eds.), *Exceptionality. Vol. 12. Exceptionality* (pp. 129–144). . https://doi.org/10.1207/s15327035ex1203_2.
- Williams, J. P., Nubla-Kung, A. M., Pollini, S., Stafford, K. B., Garcia, A., & Snyder, A. E. (2007). Teaching cause-effect text structure through social studies content to at-risk second graders. *Journal of Learning Disabilities*, 40(2), 111–120. <https://doi.org/10.1177/00222194070400020201>.
- Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge: Cambridge University Press.
- Zabrocky, K., & Ratner, H. H. (1992). Effects of passage type on comprehension monitoring and recall in good and poor readers. *Journal of Reading Behavior*, 24(3), 373–391. <https://doi.org/10.1080/10862969209547782>.