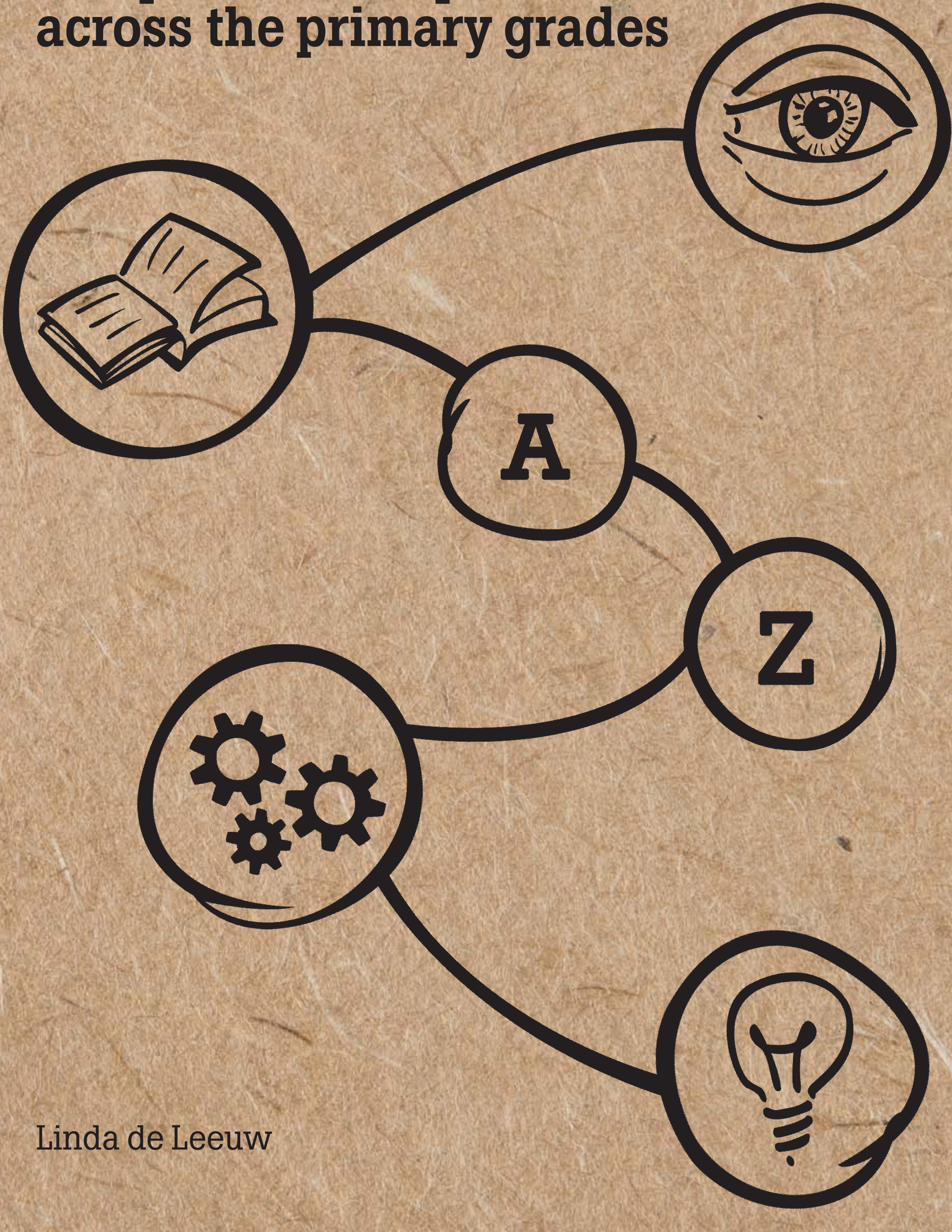
Understanding reading comprehension processes across the primary grades



Linda de Leeuw

Understanding reading comprehension processes across the primary grades

Linda Charlotte de Leeuw

Radboud University

The studies described in the present thesis were funded by the Dutch Ministry of Education, Culture, and Science and conducted at the Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

ISBN: 978-90-9029186-4

Cover lay-out
Jaap Koek

Printed by
Krex Vormgeving

Understanding reading comprehension processes across the primary grades

Proefschrift

ter verkrijging van de graad van doctor aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. Th.L.M. Engelen,
volgens besluit van het college van decanen in het openbaar te verdedigen
op dinsdag 13 oktober 2015
om 10.30 uur precies

door

Linda Charlotte de Leeuw
geboren op 30 september 1985
te Nieuwegein

Promotor

Prof. dr. L. T. W. Verhoeven

Copromotor

Dr. P. J. C. Segers

Manuscriptcommissie

Prof. dr. P. A. Coppen

Prof. dr. T. J. M. Sanders (UU)

Dr. M. Van der Schoot (VU)

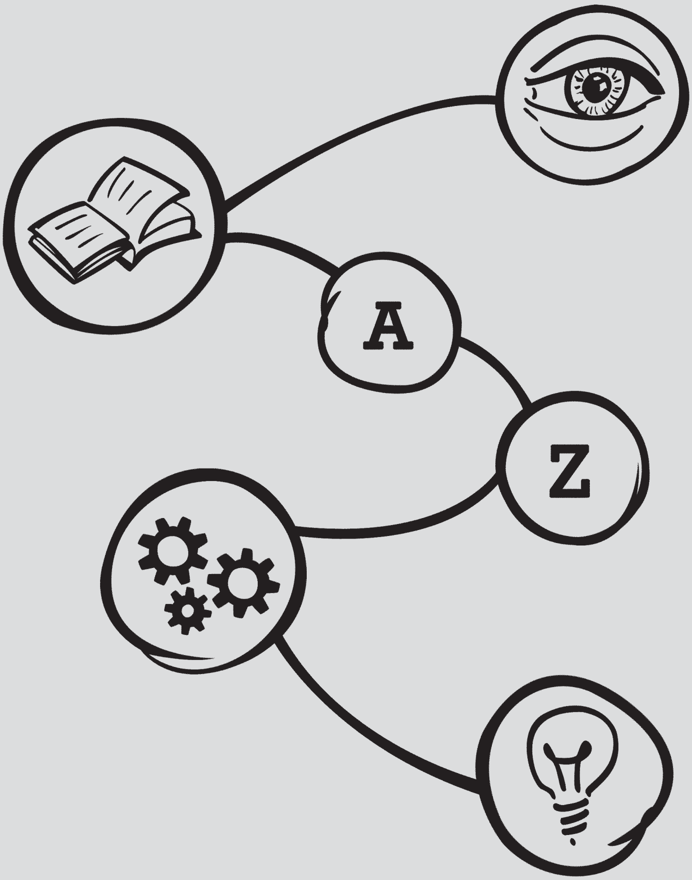
Voor mama en oma,
als dank voor deze wijze les:

Beroem je niet op komende successen,
Ook al bereik die straks voor jezelf heel graag,
Probeer er liever keihard aan te werken...
Leef vandaag!

Contents

CHAPTER 1

General introduction



Reading comprehension enables readers to acquire knowledge from a written context, which is considered a key factor in school success. The main goal of reading education is, therefore, to teach students not only how to read a text for comprehension (the *process* of reading), but also to remember the information from a text (the *product* of reading). In middle to late elementary school, the focus of reading education changes from *learning to read* to *reading to learn*. Previous research has found that both the process and product of reading are highly associated with characteristics related to the student, to the text, and to the reading task. It is therefore crucial to understand how students from 3rd to 6th grade read expository texts for comprehension to decide which texts and tasks optimize both reading comprehension processes and products for this age group. Nevertheless, few studies have been conducted that examine the real-time reading processes of developing readers. Nor have these real-time processes been related to learning from texts. The present thesis therefore aimed to gain insight into the student-related, text-related and task-related characteristics of the process and products of reading.

Reading comprehension processes

Reading comprehension can be described as the outcome of comprehension processes that occur during reading. To comprehend a text, readers must not only decode it; they must also create a representation of it. This ultimately results in a mental model that is stored in long-term memory. This section describes the most influential reading models, how the reading comprehension processes can be measured, and how the processes of reading result in a mental model after reading.

Modeling reading comprehension

Reading comprehension processes aim to build a coherent text representation. Discourse psychologists traditionally describe reading along the lines of bottom-up and top-down processes (Graesser, 2007; Kintsch, 2005). In a bottom-up approach, the reader sequentially builds a coherent representation by integrating the information of a sentence within the current representation. Top-down processes are thought to guide comprehension such as background knowledge of scripts and reading strategies.

A number of theoretical models have been proposed that aim to describe how readers construct a coherent text representation. One of the most comprehensive and influential models is the Construction-Integration model (Kinstch & Van Dijk, 1978; Kintsch, 2004). This model assumes that three different levels of text representations are built while reading. First, it is important that the reader understands the sentences within the text, which is called the parser or surface code. Second, the reader must understand how the sentences and segments cohere, leading to a coherent text-based representation. Third, the text-based representation needs to be integrated with prior knowledge, resulting in a situation model (or mental model) of the text. The quality of the text representation is determined by the depth of the representation; surface code representations are thought to be shallower than situation model representations (Kamalski, 2007).

Inference generation is important for bottom-up processes within the Construction-Integration model. An *inference* may be thought of as a connection that can or must be made to create coherence among two text segments. The Construction-Integration model distinguishes between memory-based processing and integration processing (Kinstch & Van Dijk, 1978; Kintsch, 2004). *Memory-based processes* enable readers to generate inferences by using concepts that have recently been read. These concepts are active in memory and therefore readily available for inference generation. *Integration processing* involves inference generation among text elements that need to be (re)activated. This is the case for text-based information that is no longer available in working memory, but also for related background knowledge required for integration within long-term memory. Inference processes usually occur at sentence boundaries, as evinced by several studies that show increased reading times at sentence final segments (Hirotani, Frazier, & Rayner, 2006; Rayner, Kambe & Duffy, 2000).

Top-down processes guide reading by using knowledge about scripts and reading strategies. First, background knowledge about scripts is used to generate (bridging) inferences and to solve comprehension problems that cannot be inferred from the text base (Kintsch, 2005). For example, when describing a situation in a restaurant, the roles within the script are quite strict. Usually, the customer orders and the waiter serves drinks (and not vice versa). Such knowledge may help the reader to solve comprehension problems and to understand the discourse. Second, reading strategies such as the readers' goal and level of coherence (c.f., the standard of coherence; Van den Broek, Lorch, Linderholm, & Gustafson, 2001) affect the quality of the mental model (Graesser, Singer, & Trabasso, 1994). The readers' goal in leisure reading is presumably different than it is when given the task of writing a summary or answering comprehension questions. In the latter case, the standard of coherence will be much higher. This higher standard results in extensive and better inference generation while reading.

Ultimately, both bottom-up and top-down processes require skills. Therefore, reading models should include individual variation among readers. This is especially the case when describing reading comprehension in a developmental perspective. The most influential model that focuses on reading skills is the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990; Gough, Hoover, Peterson, Cornoldi, & Oakhill, 1996). This model defines *reading comprehension* as a product of word decoding and listening comprehension. In a more recent, compatible, brain-based model, *reading comprehension* is defined as a neural network in which a memory component stores words in the mental lexicon. A unification component then combines words into meaningful sentences, and memory capacity controls the number of inferences made from context (Hagoort, 2005).

The more general reading-systems framework as described by Perfetti and Stafura (2014) can be seen as an integration of the different models just described. The framework encompasses both individual differences and reading comprehension processes and its interrelations (Figure 1). On the one hand, the model describes reading as a bottom-up process. It starts with visual information (at the left) and moves along word identification to the comprehension process (at the right). In this process, the reader sequentially builds a coherent text representation that is stored in long-term memory. On the other hand, the model includes top-down processes; general knowledge influences the situation model representation. Most importantly, this model also includes individual factors such as the linguistics and the writing system (pictured in the top box in Figure 1), word identification (middle box), and general knowledge (bottom box).

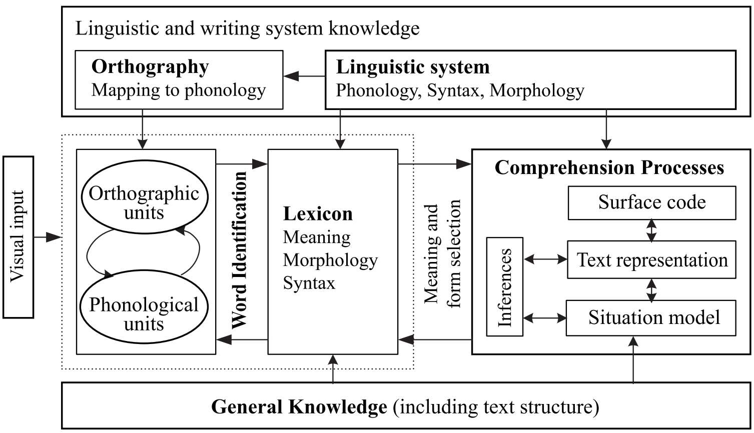


Figure 1. The components of reading comprehension from identifying words to text comprehension. Adapted from Perfetti and Stafura (2014).

Measuring real-time processes

To understand reading comprehension processes, previous studies have used several ways to measure processes while reading. First, in think-aloud protocols (Blanc, Kendeou, Van den Broek, & Brouillet, 2008) students are instructed to read a text aloud and to inform the experimenter of what they are thinking while reading. A major disadvantage of this setup is that it disrupts the reading process. In addition, children are often unable to properly vocalize their thinking because they lack metacognitive skills (Kuhn, 2000). Another method is self-paced reading (Aaronson & Scarborough, 1976): segments of the text (usually a word or sentence) are sequentially presented to the reader. Whenever the reader has finished reading a segment, he or she presses a button to receive the next one. A major downfall of this method is in its ecological validity: pressing buttons while reading interferes with the reading processes. To overcome these problems eye movements can be studied. This setup is more frequently used while examining real-time reading processes (Blythe & Joseph, 2011). The increase in the amount of eye tracking studies is due mainly to the availability of more child-friendly and less intrusive eye tracking equipment. In addition, eye trackers have become more mobile, which makes it possible to conduct eye movement studies at such locations as schools, thereby enabling large-scale eye movement studies in children.

In eye tracking research, movements of the eyes are measured by using infrared light that localizes the pupil. The frequency at which these gaze locations are generated is determined by the Hz-frequency of the eye tracking equipment. A 120 Hz eye tracker determines the position of the eye every 8 ms, whereas a 1000 Hz eye tracker provides gaze points each millisecond. To map the location of the eye to a specific position on the screen, a calibration procedure is required prior to testing. During this procedure, the participant needs to follow a dot that moves along the screen. The dot stops at several positions, usually six or nine. The eye tracking system links the position of the pupil to a specific stop. With this information, the system is able to calculate the location of the eyes on the screen. Information about gaze locations is then used to calculate fixations and saccades. *Fixations* are defined as positions at which the eye stops for at least 80 ms, which is the minimum amount of time needed for information processing. Information is presumed to be processed at these locations. *Saccades* are the movements of the eyes from one fixation to the next. Saccadic movements can be made forward (progressive) or backwards (regressive).

Fixations serve as a basis for different eye movement measures. In reading research, several measures are used, which can be subdivided into probability and durational measures. To understand the probability measures, consider reading a single sentence. You might read all of the words, but most likely you will skip some. This is reflected by *skipping probability*; the chance of skipping a word. When you continue reading, you will most often read from left to right (in western languages). But when you encounter a difficulty, you might reread previous parts of the text to solve this coherence problem. When you go back, this is referred to as a regression. *Regression probability* reflects the chance that a reader will look back to previous text segments.

Durational measures are depicted in milliseconds for a specific target word. The most common measures are gaze and regression path duration (Rayner, 1998). *Gaze duration* is the time a reader fixates on a word when encountering it for the first time, before progressing or regressing to another region. When readers skip a word, no gaze duration is calculated. *Regression path duration* can be subdivided into look back and second pass duration. *Look back duration* is the sum of all fixations on previous text. *Second pass duration* is the sum of all fixations on the target words, whenever it is reread after a regression. These latter durations reflect the time a reader spends on solving a comprehension problem.

From process to product

Both bottom-up processes and top-down processes are not only related to reading processes; they also affect the text representation that is stored in memory (Ericsson & Kintsch, 1995). The idea is that the mental model is a “network of propositions” (Kintch, 1994: 295) that improves when the number of propositions and interconnections between propositions increases. This is validated by several studies which show that more inferences lead to superior recall (Van den Broek, Rapp, & Kendeou, 2005). Nevertheless, the quality of inferences is important too (Linderholm, Virtue, Tzeng, & Van den Broek, 2004; Tarchi, 2010). This quality depends on the distance between two propositions; inferences that are drawn locally construct shallow text representations, whereas global inferences, which are drawn across larger text segments, construct deeper text representations (Graesser et al., 1994). Also, integration with background knowledge, referred to as *elaborate inference*, is considered to be more beneficial for overall learning than more text-based inferences (Graesser et al., 1994; Kalamski, 2007; Kinstch, 2004).

However, the process of reading is not *necessarily* related to the quality of the mental model. First, not all of the information that is included in the mental model during reading is necessarily remembered after reading (Just & Carpenter, 1980). This could be caused by the structure of the text. Some propositions are linked more directly to the main theme than others. As it turns out, these more directly linked propositions are recalled better after reading (Van den Broek, Young, Tzeng, & Linderholm, 1999; Van den Broek, Helder, & Van Leijenhorst, 2013). Second, less skilled readers might use compensational strategy behavior (Walczyk, 2000), such as slowing down, looking back, pausing or shifting their attention (Perfetti, 1988). By compensating for their low skills, these readers overcome reading problems and may end up with good mental models. However, not all less skilled readers will increase the amount of cognitive energy to increase comprehension. As a result, reading comprehension may not be linearly related to comprehension outcomes.

Variation in reading comprehension

Reading comprehension is affected by student-related, text-related and task-related characteristics. Individual variation among readers affects both the process and product of reading comprehension. Skills that are found to be related to reading comprehension include both linguistic and cognitive skills. Text characteristics such as word type, text difficulty, and text length can shape reading comprehension processes. Finally, reading tasks provided during text processing can help the reader to construct a coherent model.

Student-related characteristics

Reading comprehension processes vary widely between readers. In adult readers, the processes of skilled and non skilled readers are different. More proficient readers skip more words (Roy-Charland, Saint-Aubin, Klein, & Lawrence, 2007) and have shorter gaze durations (for an overview see Radach & Kennedy, 2013). Also for developing readers, there is ample evidence that the processes of skilled and less skilled readers differ (Blythe & Joseph, 2011, Van der Schoot, Reijntjes, & Van Lieshout, 2012). Finally, differences between children and adults are found; when reading a similar text, previous text segments are read more often by younger developing readers (20-25% of the time) than by more proficient readers (10-15%) (Rayner, 1985; Reichle, Rayner, & Pollatsek, 2003). As student-related and text-related characteristics were not considered when comparing these groups, it remains unclear whether differences between children and adults are due to age, skill, or an interrelation of the two factors.

The product of reading is influenced by individual variation in both the linguistic and the cognitive domain. Within the linguistic domain, previous research has shown several different skills to be important, including decoding (Huestegge, Radach, Corbic, & Huestegge, 2009; Verhoeven & Perfetti, 2008), vocabulary (Calvo, Estevez, & Dowens, 2003; Singer, Andrusiak, Reisdorf, & Black, 1992), and reading comprehension skills (McMaster, Espin, & Van den Broek, 2014). Note that Perfetti's and Stufura's 2014 model includes all of these skills.

Within the cognitive domain, memory is also found to be important for reading comprehension, as all "processes take place within a cognitive system that has pathways between perceptual and long-term memory and limited processing resources" (Perfetti & Stafura, 2014: 25). Research on inference generation supports this view by showing that the quality of the mental model is highly related to the number of inferences that are generated during reading (Linderholm et al., 2004). In particular, this is the case because developing readers' working memory might be overloaded with lower-level processing (i.e., decoding, vocabulary) during text reading. This might limit the working memory capacity available for higher-level text processing (Just & Carpenter, 1992) such as text integration, thereby producing a qualitatively inferior mental model. Moreover, previous research has found a relation between short-term memory and working memory and reading comprehension (Cain, Oakhill, Barnes, & Bryant, 2001; Cain, Oakhill, & Bryant, 2004; Daneman & Merikle, 1996), confirming the contribution of these cognitive skills to reading.

Text-related characteristics

Text-related characteristics also influence the reading comprehension processes. Two characteristics can be considered: text complexity and text length. Whenever the text is more complex, reading is slowed in adults (Hyönä, 2011; Clifton & Staub, 2011; Rayner, Chace, Slattery, & Ashby, 2006). But this is especially true for younger and less skilled readers (Häikiö, Bertram, Hyönä, & Niemi, 2009; Rayner, 1986). Text difficulty is determined by factors such as word length and word frequency, which are often found to influence the reading processes of both adults and children (Just & Carpenter, 1980; Benjamin, 2012). Furthermore, word class and the position of a word within a sentence also influence reading, with function words being skipped more often (Roy-Charland et al., 2007) and sentence final words showing sentence wrap-up effects (Hirotani et al., 2006; Rayner et al., 2000).

Another text characteristic is the length of text. Multiple-paragraph texts require the reader to adapt reading processes throughout the text. Previous research shows that reading processes become faster at the end of a text (Linderholm et al., 2004). This could be due to the fact that processing is more efficient (Bell, 2011, Linderholm et al., 2004), or to reader fatigue (Graesser et al., 1994; Van den Broek, Risden, & Husebye-Hartman, 1995) or to mind wandering (Nguyen, Binder, Nemier, & Ardoin, 2014). The effect of the first would not (or might even positively) affect reading comprehension, whereas the latter two would negatively affect reading comprehension.

Task-related characteristics

Reading comprehension tasks are often used in educational settings to enhance learning outcomes: e.g., cloze tasks, inference questions, and summary writing. When performing a task, the reader is encouraged to interact with the text. However, not all assignments are found to improve learning outcomes. In line with the Construction Integration model, a well-designed task enhances the number and the quality of inferences that readers make (Linderholm et al., 2004; Van den Broek et al., 2001). When more inferences are generated, this leads to a more interconnected network of propositions. And propositions that have more connections are better recalled. Hence, the task should aid the reader to actively make inferences.

Furthermore, the quality of the inferences is also important. Local (more surface code-based) inferences are presumed to lead to shallower presentations. Global (more text-based) inferences connect two or more sentences and are qualitatively superior to local inferences. Nevertheless, memory for text is best when the text is integrated with prior knowledge (elaborate inferences). A task that enhances the generation of more and higher-level inferences is therefore presumed to be better for learning (Cerdán, Vidan-Abarca, Martínez, Gilabert, & Gil, 2009; Wixon, 1983), though it is unclear whether different tasks elicit similar of different effects among readers. For example, higher-level tasks may be very effective for skilled readers, but they may overload the memories of less skilled readers' and so lead to poorer learning results.

The present thesis

The above overview of the literature shows that reading skills are related both to the process and to the product of reading. However, few studies have considered this phenomenon in a developmental perspective. For this reason, the main focus of the present thesis is on individual variation in reading processes of students across the primary grades. In particular, the reading processes of children in Grade 3-6 are studied, because, in general, these readers have finished learning to read and now read to learn. This thesis also focuses on the effects of text-related and task-related characteristics. Text-related characteristics such as word type, text difficulty, and text length are found to influence text processing; but it remains unclear how these factors affect reading in a developmental perspective. Moreover, including text-related and task-related characteristics makes it possible to examine not only inter-individual but also intra-individual variation in reading comprehension processes. Finally, the combination of reading processes, products, individual variation and examined interrelations among them has not been considered in previous research. Therefore, the main aim of the research presented in this thesis is to develop further understanding of text comprehension processes by considering how students-related, text-related and task-related characteristics influence the process and product of reading.

The present thesis describes four studies in which these research questions were addressed. *Chapter 2* starts by examining the real-time processes of 24 third-grade and 20 fifth-grade students. All students were asked to read both a relatively easy text (i.e., one below their grade level) and a more difficult text (i.e., one at their grade level). First, individual differences with respect to word decoding, reading comprehension, short-term memory and working memory were taken into account. Second, text characteristics related to the difficulty of the text were examined.

In *Chapter 3*, the effect of real-time reading process on the relation between student-related characteristics and text comprehension are examined in 4^{th} graders. Students' eye movements were recorded as they read four expository texts and subsequently answered text comprehension questions. Children's reading processes were examined for the heading, first sentence, and final sentence to determine both differences in reading strategy behavior and sentence wrap-up effects.

Chapter 4 examines the real-time reading processes of 6^{th} grade students as they read expository texts consisting of one introductory paragraph and three sections that were each three paragraphs long. All paragraphs started with a heading. The main aim was to determine the time course of effects of comprehension processes during and after reading, including text-related effects of section and paragraph, and to determine the role of student-related characteristics (word decoding, vocabulary, comprehension skill, short-term memory, working memory, and non-verbal intelligence). Seventy-three sixth graders read two texts and subsequently performed two text-comprehension tasks: i.e., they answered multiple-choice questions and performed a related-judgment task that measures knowledge representations. Eye movements were recorded and total reading times of the heading and remainder of the paragraph were analyzed.

The effects of different reading comprehension tasks in 5th grade are examined in *Chapter 5*. The tasks were designed to stimulate reading comprehension at different levels. The first task was a gap filling task that focused on surface code processes. The second task involved inference questions, which are at the level of the text base. The final task was a summary writing task, which manifests at the level of the situation model. Students practiced with one of the tasks for three weeks, after which the effect of this practice on incidental word learning was tested using a vocabulary interview. The study examined the effects of the different tasks. Interactions with skills and capabilities of the students - such as general vocabulary knowledge and working memory - were also determined.

Finally, a general discussion is provided. Chapter 6 reviews and discusses the results of the four experiments described in this thesis and provides an overview of its contribution to current theories on reading comprehension. Furthermore, limitations and suggestions for future research and a general conclusion and practical implications are presented.

References

Aaronson, D., & Scarborough, H. S. (1976). Performance theories for sentence coding: some quantitative evidence. *Journal of Experimental Psychology: Human Perception and Performance, 2*(1), 56-70. doi: 10.1016/S0022-5371(77)80052-2.

Bell, A. (2011). Re-constructing Babel: Discourse analysis, hermeneutics and the Interpretive Arc. *Discourse Studies, 13*(5), 519–568. doi:10.1177/1461445611412699.

Benjamin, R. G. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review, 24*(1), 63-88. doi: 10.1007/s10648-011-9181-8.

Blanc, N., Kendeou, P., Van den Broek, P. W., & Brouillet, D. (2008). Updating situation models during reading of news reports: Evidence from empirical data and simulations. *Discourse Processes, 45*(2), 103-121. doi: 10.1080/01638530701792784.

Blythe, H. I., & Joseph, H. S. S. L. (2011).Children's eye movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye-movements*. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199539789.013.0036.

Cain, K., Oakhill, J., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859. doi: 10.3758/BF03196414.

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

Calvo, M. G., Estevez, A., & Dowens, M. G. (2003). Time course of elaborative inferences in reading as a function of prior vocabulary knowledge. *Learning and Instruction*, *13*, 611–631. doi: 10.1016/S0959-4752(02)00055-5.

Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction*, *19,* 13-27. doi:10.1016/j.learninstruc.2007.12.003.

Clifton, C., Jr., & Staub, A. (2011). Syntactic influences on eye movements during reading. In S. Liversedge, Iain Gilchrist, and Stefan Everling (Eds), *Oxford Handbook on Eye-movements. Chapter 49* (pp. 895-909). Oxford: Oxford University Press.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4). 422-433. doi:10.3758/s13423-013-0565-2.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological review*, *102*(2), 211-245. doi: 10.1037/0033-295X.102.2.211.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*, 6-10. doi: 10.1177/074193258600700104.

Gough, P. B., Hoover, W. A., Peterson, Cornoldi, C., & Oakhill, J. (1996). Some observations on a simple view of reading. In C. Cornoldi, & J. Oakhill (Eds). *Reading comprehension difficulties: Processes and intervention* (pp.1-13). New Yersey: Mahwah, Lawrence Erlbaum Associates.

Graesser, A. (2007). An introduction to strategic reading comprehension. In D.S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 465-496). Mahwah, NJ: Erlbaum.

Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371-395. doi:10.1037/0033-295X.101.3.371

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in cognitive sciences*, *9*(9), 416-423. doi:10.1016/j.tics.2005.07.004.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology, 102,* 167-181. doi:10.1016/j.jecp.2008.04.002.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language, 54*(3), 425-443. doi:10.1016/j.jml.2005.12.001.

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

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. doi: http://dx.doi.org/10.1016/j.visres.2009.09.012.

Hyönä, J. (2011). Foveal and parafoveal processing during reading. In S. Liversedge, I. Gilchrist, & S. Everling (Eds.) *The Oxford Handbook of Eye-Movements* (pp. 819-838).Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199539789.013.0045.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4). 329-354. doi: http://dx.doi.org/10.1037/0033-295X.87.4.329.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122-149. doi: 10.1037/0033-295X.99.1.122.

Kamalski, J. (2007). *Coherence marking, comprehension and persuasion: On the processing and representation of discourse*, Utrecht: LOT.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294-303.

Kintsch, W. (2004). The Construction-Integration model of text comprehension and its implications for instruction. In: R.B. Ruddell, & N. J. Unrau, (Eds.), *Theoretical Models and Processes of Reading*. *Fifth Edition* (pp.1270-1328). Newark, USA: International Reading Association.

Kintsch, W. (2005). An Overview of Top-Down and Bottom-Up Effects in Comprehension: The CI Perspective. Discourse Processes, 39(2-3), 125-128. doi: 10.1080/0163853X.2005.9651676.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394. doi:10.1037/0033-295X.85.5.363.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*(5), 178-181. doi:10.1111/1467-8721.00088.

Linderholm, T., Virtue, S., Tzeng, Y., & Van den Broek, P. (2004). Fluctuations in the availability of information during reading: Capturing cognitive processes using the Landscape Model. *Discourse Processes, 37,* 165-186. doi:10.1207/s15326950dp3702_5.

McMaster, K. L., Espin, C. A., & Van den Broek, P. (2014). Making connections: Linking cognitive psychology and intervention research to improve comprehension of struggling readers. *Learning Disabilities Research & Practice*, *29*(1), 17–24. doi: 10.1111/ldrp.1202.6

Nguyen, K., Binder, K., Nemier, C., & Ardoin, S. (2014). Gotcha! Catching kids during mindless reading, *Scientific Studies of Reading*, *18*(4), 274-290. doi:10.1080/10888438.2013.876424.

Perfetti, C. A. (1988). Verbal efficiency in reading ability. In G. E. MacKinnon, T. G. Waller, & M. Daneman (Eds.), *Reading research: Advances in theory and practice, Vol. 6* (pp. 109-143). New York: Academic Press, Inc.

Perfetti, C. A., & Stafura, J. (2014) Word knowledge in a theory of reading comprehension, *Scientific Studies of Reading, 18*, 22–37. doi:10.1080/10888438.2013.827687.

Radach, R., & Kennedy, A. (2013). Eye-movements in reading: Some theoretical context. *The Quarterly Journal of Experimental Psychology*, *66*(3), 429-452. doi:10.1080/17470218. 2012. 750676.

Rayner, K. (1985). The role of eye movements in learning to read and reading disability. *Remedial and Special Education, 6*(6), 53-60. doi: 10.1177/074193258500600609.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*, 211-236.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Pychological Bulletin, 124*(3), 372-422. doi: 10.1037/0033-2909.124.3.372.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye-movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241-255. doi:10.1207 /s1532799xssr1003_3.

Rayner, K., Kambe, G., & Duffy, S. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology, 53A*(4), 1061-1079. doi: 10.1080/713755934.

Roy-Charland, A., Saint-Aubin, J., Klein, R. M., & Lawrence, M. (2007). Eye-movements as direct tests of the GO model for the missing-letter effect. *Perception & Psychophysics, 69*(3), 324-337. doi: 10.3758/BF03193753.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*, 445-526. doi: 10.1017/S0140525X03000104.

Singer, M., Andrusiak, P., Reisdorf , P., & Block, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition*, *20*(5), 539-548.

Tarchi, C. (2010). Reading comprehension of informative texts in secondary school: A focus on direct and indirect effects of reader's prior knowledge. *Learning and Individual Differences*, *20*, 415–420. doi:10.1016/j.lindif.2010.04.002.

Van den Broek, P. W., Helder, A., & Van Leijenhorst, L. (2013). Sensitivity to structural centrality: Developmental and individual differences in reading comprehension skills. In M. A. Britt, S. R. Goldman, & J. F. Rouet (Eds.),?*Reading: From Words to Multiple Texts* (pp. 132-146). New York: Routledge.

Van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition*, *29*(8), 1081-1087. doi: 10.3758/BF03206376.

Van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist approaches in accounts of reading comprehension. *Discourse Processes, 39,* 299-316. doi: 10.1080/0163853X.2005.9651685.

Van den Broek, P. W., & Risden, K., & Husebye-Hartman, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In: Lorch, Jr., R. F. & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353-373). Hillsdale, NJ: Erlbaum.

Van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading. In H. Van Oostendorp & S. R. Goldman (Eds.),?*The construction of mental representations during reading*?(pp. 71-98). Mahwah, NJ: Erlbaum.

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. doi:10.1007/s11145-011-9337-4.

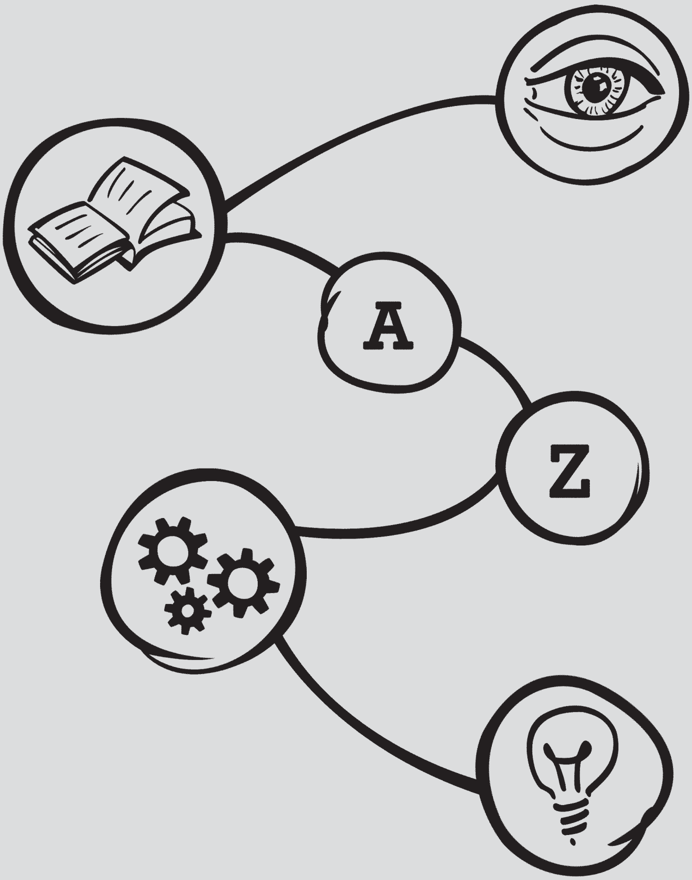
Verhoeven, L., & Perfetti, C. (2008). Advances in text comprehension: Model, process and development. *Applied Cognitive Psychology, 22*, 293–301. doi:10.1002/acp.1417.

Walczyk, J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly,* 35(4), 554-566. doi:http://www.jstor.org/stable/748099.

Wixon, K. (1983). Postreading question-answer interactions and children's learning from text. *Journal of Educational Psychology*, *30*(3), 413-423. doi: 10.1037/0022-0663.75.3.413.

CHAPTER 2

Role of text and student characteristics in real-time reading processes across the primary grades[1]



---

[1]This paper has been accepted for publication:
Leeuw, L. de, Segers, E., & Verhoeven, L. (in press). Role of Text and Student Characteristics in Real-time Reading Processes across the Primary Grades. *Journal of Research in Reading*. doi:10.1111/1467-9817.12054

Abstract

Although much is known about beginning readers using behavioural measures, real-time processes are still less clear. The present study examined eye movements (skipping rate, gaze, look back, and second pass duration) as a function of text-related (difficulty, and word class) and student-related characteristics (word decoding, reading comprehension, short-term and working memory). Twenty-four third and 20 fifth graders read a relatively easy (below grade level) and more difficult text (at grade level). The results showed that skipping rate mainly relied on text characteristics and a three-way interaction of grade, text difficulty, and word class. Gaze durations depended mostly on student characteristics. Results on look backs showed more and longer look backs in difficult texts. Finally, second pass duration mostly relied on grade level. To conclude, this study shows that both student and text characteristics should be taken into account when studying online text reading development.

Introduction

As eye trackers become more and more child friendly, research studying children's eye movements in reading is increasing. Several studies showed that eye movement patterns of beginning readers are different from those of adults (for a review see Blythe & Joseph, 2011). And although differences between students across grades have been found in eye movement control, as evidenced by studies on binocular coordination (Blythe, Liversedge, Joseph, White, Findlay, & Rayner, 2006) and parafoveal processing (Häikiö, Betram, & Hyönä, 2010; Häikiö, Betram, Hyönä, & Niemi, 2009), these oculomotor effects did not show an effect on reading development (Huestegge, Radach, Corbic, & Huestegge, 2009; Rayner, 1986) and are more likely to be associated with difficulties readers encounter (Hyönä & Olson, 1995).

With respect to text processing, it has been suggested that eye movements reflect processing activities associated with reading comprehension (Rayner, 1985; Rayner, Chace, Slattery, & Ashby, 2006; Rayner, Juhasz, & Pollatsek, 2005; Rayner & Liversedge, 2011); whenever readers encounter a difficulty in the text, reading is slowed down resulting in more and longer fixations and more regression to previous text segments (Rayner & Slattery, 2009). The problem with this account is that effects can be caused by text characteristics (Hyönä, 2011), but also by reading skill (McConkie, Zola, Grimes, Kerr, Bryant, & Wolff, 1991) or age (Blythe & Joseph, 2011). Previous eye tracking studies have found that text characteristics, such as word class (Roy-Charland, Saint-Aubin, Klein, & Lawrence, 2007; Blythe, Liversedge, Joseph, White, & Rayner, 2009) and text difficulty (Rayner et al., 2006) influence text processing. Also, studies on adults and adolescents found that text reading difficulties can be associated with reading proficiency reflected by decoding and comprehension skills (Kuperman & Van Dyke, 2011) and cognitive abilities such as short-term memory (De Abrue, Gathercole, & Martin, 2011) and working memory (Nation, 2007). Previous research has not been successful in disentangling the effect of grade level, cognitive skills and reading skills on real-time processing (Blythe & Joseph, 2011) whereas such studies including these measures can be seen as highly informative in explaining individual differences in reading ability and the time course of these effects. In the present study, we therefore examined the eye movements of developing readers at different grade levels (third and fifth grade) when reading an easy and a more difficult text as a function of word class, children's reading proficiency, short-term and working memory.

Text-related characteristics

Text comprehension has been found to be influenced by many factors that increase the complexity of the text (McNamara, Kintsch, Songer, & Kintsch, 1996). Therefore, readability formulas used to determine text difficulty generally include measures of word length, word frequency, sentence length and the percentage of familiar words (Benjamin, 2012). Word length and word frequency are highly related and longer and less frequent words are less easy to process (Just & Carpenter, 1980). Also, longer sentences place a higher demand on working memory, which increase the difficulty (De Abrue et al., 2011). Finally, also the density of known words (Vermeer, 2000) and content and function words (Graesser, McNamara, Louwerse, & Cai, 2004) were found to influence text difficulty.

Although there is only limited research evidence, it is generally assumed that the *overall complexity* of the text has an impact on children's eye movements during reading (Blythe et al., 2009; Chamberland, Saint-Aubin, & Légère, 2013). This assumption is based on evidence from studies focussing on one aspect of text difficulty, such as word frequency, age of acquisition, word length and predictability, and grammatical complexity influence eye movement patterns (for an overview see Hyönä, 2011; Clifton & Staub, 2011). When encountering such difficulties, readers tend to focus on particular text elements for a longer period of time, slowing down their foveal and parafoveal processing (Henderson & Ferreira, 1990). This results in slower reading times of adult skilled readers, but also, or even more so, for young and less skilled readers (Häikiö et al., 2009; Rayner, 1986). Other evidence shows interpersonal differences among easy and difficult texts. Pirozollo and Rayner (1978, as cited in Rayner, 1985) showed dyslexic students show similar eye movement patterns for dyslexic and reading-matched controls when reading materials were adapted to their reading level, but distinctive patterns when reading a text that is more difficult appropriate for their age. Similar results were found for adults (Rayner et al., 2006), indicating that eye movements not only depend on the skills of a reader, but also on the difficulty of the text (Oakland & Lane, 2004).

Previous studies that focus on individual effects of text characteristics have found these characteristics to be important at different stages of processing. Very robust affects that influence very early reading processes reflected by first fixation duration include word length effects (Joseph, Liversedge, Blythe, White, & Rayner, 2009) which are found to be similar for mono spaced and relative fonts (Hautala, Hyönä & Arco, 2011). In addition, several studies have shown word frequency to be important in first fixation durations (Blythe et al., 2009; Joseph, Nation, & Liversedge, 2013). Moreover, word length and frequency effects have been found to be larger for children compared to adult readers (Joseph et al., 2009), though no difference is found between skilled and less skilled readers (Hyönä & Olson, 1995). When considering effects of higher order processes such as syntactic complexity (Joseph & Liversedge, 2013) and pragmatic coherence (Joseph, Liversedge, Blythe, White, Gathercole, & Rayner, 2008; Vauras, Hyönä, & Niemi, 1992), similar affects are found for adults and children, although the time-course of the effects was found to be delayed for children.

Word class is another text-related characteristic that appears to influence eye movements. Words classes can be subdivided in function and content words (Fromkin, 2000; Chamberland et al., 2013). Function words are mostly grammatical in nature and express grammatical relationships between lexical entities in the sentences. It is a closed-class of words and includes a fixed set of, for example, prepositions, determiners, and auxiliaries. These words are often short and frequent. Content words constitute an open-class. For example, adding pre- or suffixes generates new words that can be adjoined to the group of content words. This class includes lexical words such as nouns, adjectives, verbs, and adverbs.

Function and content words are processed differently during reading, with function words being skipped more often (Roy-Charland et al., 2007). Carpenter and Just (1983) found that 83% of the content words and only 38% of the function words were fixated. This could be due to the nature of function words, in the sense that they tend to be much more frequent, predictable and shorter than content words. When controlling for each of those factors, however, Chamberland, et al. (2013) still reported similar effects, albeit smaller (66% of the content words were fixated compared to 57% of the function words). In sum, these results show within reader variability in eye movements as a function of text difficulty.

Student-related characteristics

Online reading processes also depend on individual cognitive and reading abilities (Blythe & Joseph, 2011). During the primary school years, children become faster in word decoding every year (Verhoeven & Van Leeuwe, 2008). And with regard to reading comprehension, skilled readers more easily draw inferences and build more elaborate mental models of the text (McNamara & O'Reilly, 2009). Memory capacity is related to both decoding skills as well as reading comprehension (Kintsch, 2004).

Beginning readers fixate on words more often than more proficient readers (Rayner, 1985) and adult readers (Lester, Nagle, Johnson, & Fisher, 1979; McConkie et al, 1991). Both the number and duration of fixations appear to decrease with age and proficiency (for an overview see Radach & Kennedy, 2013). More and longer fixations reflect processes beginning readers are particularly dealing with since their decoding lacks fluency (Verhoeven & Van Leeuwe, 2008). In particular, students learning to read in an orthographically shallow language, such as Dutch, may benefit from increased automated decoding skills since their parafoveal view will accordingly increase as well (Häikiö et al., 2009).

Look back patterns are also different in beginning as compared to more proficient readers. Looking back to previous text segments has been found to indicate processing problems; the reader encounters a problem integrating the text into the previously read segment (i.e., a comprehension problem). When reading a similar text, previous text segments are read more often by beginning readers (20-25% of the time) than more proficient readers (10-15%) (Rayner, 1985; Reichle, Rayner, & Pollatsek, 2003).

Differences in reading skill may not only lead to faster reading times, but also to different reading patterns. Skilled readers tend to pay more attention to important words than less important words (Kaakinen, Hyönä, & Keenan, 2003; Reynolds, 2000; Van der Schoot, Vasbinder, Horsley, & Van Lieshout, 2008) and spent more time on mental model updating (Schroeder, 2011). Moreover, Van der Schoot et al. (2008) found that less skilled readers do not invest more processing time in important text elements. Skilled readers, on the other hand, spend more time looking back to previous text segments when they encounter an important word. This extra processing time is considered as time invested in the integration of important text elements into the mental model. More proficient readers also use specific skills that enable them to read difficult words and sentences. Examples are metacognitive knowledge and knowledge about reading strategies (McNamara & O'Reilly, 2009), resulting in differences in eye movement patterns. In addition, skilled readers are better at monitoring their comprehension which may result in more regressive eye movements compared to less skilled readers (Oakhill & Cain, 2007; Van der Schoot, Reijntjes, & Van Lieshout, 2012). Although reading skills generally develop as a function of grade level, older readers are not necessarily better readers. Poor readers in 5th grade tend to have longer gaze durations than good readers in 3rd grade (Lester et al., 1979).

Memory is an important cognitive factor that needs to be taken into account when studying online reading processes (Swanson & Ashbaker, 2000). There is empirical evidence that comprehension of children with poor short-term and working memory is relatively weak (Nation, 2007; Swanson & Ashbaker, 2000), although working memory is found to be a more important predictor than short-term memory (Daneman & Merikle, 1996). Poor readers are more involved in lower-level text processing, which limits the amount of working memory capacity available for higher-level text processing (Just & Carpenter, 1992). In addition, poor readers are most often slower readers. And, when processing demands increase by for example a reading aloud task, their reading slows down relatively much compared to good readers (Vorstius, Radach, & Lonigan, 2014).

Although no eye-tracking studies focused on the relation of short-term memory and reading comprehension, various studies have found indications that short-term memory influences reading comprehension (Molfese, Molfese, & Modgline, 2001) and are related to vocabulary knowledge and syntactic processing in particular (De Abrue et al., 2011). In addition, there is ample evidence suggesting that working memory capacity is associated with eye movements during reading (Kaakinen, Hyönä, & Keenan, 2002; Kaakinen et al., 2003). In their studies, Kaakinen et al. (2002; 2003) found adult readers with high working memory capacity allocate their attention to relevant information better at both the gaze and look back of relevant regions. Readers with low working memory capacity also allocate their attention to relevant information, but do so by looking back at the relevant information and not by spending more time on processing in gaze duration. These findings suggest that good readers are better at detecting important information for the mental model when they first encounter this information, and are thus faster at constructing their mental model.

The present study

Previous research has shown that online reading processes can be seen as a function of student- and text-related characteristics. However, a developmental perspective on online reading processes is generally lacking. As more difficult texts slow down reading of skilled readers, the question arises whether differences in eye movements are driven by reading skill or age, and whether the effect is confounded by text difficulty. Stanovich (1986) argued that reading patterns are also determined by the level of the text, and not only by the proficiency of the reader. On the other hand, Blythe and Joseph (2011) showed age related effects are similar for studies controlling for text difficulty and studies using non-age appropriate materials, suggesting that developmental changes are not affected by text difficulty. To date, the confounding role of text difficulty on eye movements in children remains unclear, because no research thus far has combined text difficulty, grade level and reading skill in one design.

With age, readers are becoming more proficient readers. Hence, a similar text is easier to read and therefore results in different text processing reflected by differences in eye movements. Most studies discussing developmental changes focused on averaged eye movement scores, not taking into account individual differences in skill or text difficulty. For this reason, it remains unclear to what extent the developmental changes found in previous studies are caused by subskills involved in reading processes, or whether these differences are only age-related. And, although short-term memory and working memory are found to be related to reading comprehension, few studies have investigated their relation with online measures in developing readers.

To sum up, the aim of the present study was to gain more insight into the development of eye movements and the role of reading skill, working memory and text difficulty by comparing eye movements of readers of Grade 3 and 5. A cross-sectional eye-tracking study was conducted in which children read an easy (below grade level) and a more difficult text (at grade level). Reading times of content words (more important for text understanding) were compared to those of function words (less important for text understanding). The following research questions were addressed:

To what extent do eye movements of Grade 3 and 5 students differ as a function of text characteristics (i.e., text difficulty and word class)?

To what extent do student characteristics (i.e., word decoding, short-term memory, working memory, and reading comprehension,) contribute to the variation in eye movements?

With respect to the first question, we hypothesized that eye movements are predicted by the text-related characteristics. Whenever a text is more difficult, we expected less skipping, longer reading times and more and longer look backs in particular. Furthermore, differences among grades were also expected, since it can be assumed that monitoring skills are more apparent in Grade 5. Therefore, Grade 3 students are expected to show relatively fewer regressions in the more difficult text, whereas Grade 5 students are expected to look back more often when reading a more difficult text. Finally, we expect function words to be skipped more often, show shorter gaze, look back and second pass durations. Finally, we expect 5th graders to be more consistent in skipping function words, since these students are more experienced readers. Third graders are expected to be less experienced and hence slow down reading whenever reading is difficult, resulting into longer gaze durations and less skipping. With respect to regressive eye movements, we expect Grade 5 students to be applying monitoring skills more often, especially in difficult texts.

With respect to the second question, we expected all student-related characteristics to predict eye movement patterns; reading times were expected to be shorter for students in higher grades, with assumable higher levels of decoding, reading comprehension skills and memory capacity. Lower levels skills such as short-term memory and decoding are expected to show effects for gaze durations in particular, whereas higher level skills such as working memory and reading comprehension are expected to influence look back and second pass duration. Furthermore, grade and skills were expected to show an interaction, because building a coherent text representation (i.e. mental model) is assumed to be most successful when readers have both the experience to link text segments *and* the memory capacity available to store information that can be linked.

Method

Participants

Students from two Dutch primary schools participated: two 3rd grade and two 5th grade classes. From the 84 students, some were excluded from analyses, because they were diagnosed with dyslexia (*n* = 9) or had reading comprehension scores that were more than two standard deviations from the mean (*n* = 2). Also, participants (*n* = 29) were removed from data analysis due to unusable fixation data caused by children's movements after calibration, which is normal in eye-tracking settings without a chin rest (Navab, Gillespie-Lynch, Johnson, Sigman, & Hutman, 2012). In total, 24 third grade students (12 girls, 12 boys, *M*age = 8 years11 months, age range from 7 years 8 months until 10 years 2 months) and 20 fifth grade students (13 girls, 7 boys, M_{age} = 10 years 10 months, age range from 9 years 11 months until 12 years) were included in the analyses. Participants had a normal non-verbal IQ, all scoring above the 25th percentile (*Standard Progressive Matrices;* Raven, 1960). Grade 3 students (M = 36.64, SD = 5.35) did differ with respect to non-verbal IQ from Grade 5, M = 40.50, SD = 4.63), t (46) = 2.70, p = .009, d = -0.77. However, non-verbal IQ was not found to predict eye movements in any form and is hence not included as a predictor in the present study.

Materials

Short-term memory (STM). STM was measured using a forward digit span memory task (*WISC-III^{NL}*, Kort et al., 2005). The researcher read aloud a string of digits using a falling intonation and pausing one second between the digits. The students were instructed to remember the digits in the same order. The strings started short (two digits) with two attempts for each string length. Whenever children correctly remembered at least one of two strings, the researcher continued with a longer string, adding one digit until a maximum (nine digits) was reached. Each correctly remembered string accounted for one point with a maximum of 16 and were included in the analyses as z-scores.

Working memory (WM). WM was measured by a backward digit span memory task (*WISC-III^{NL}*, Kort et al., 2005). This task is similar to the STM task, however, students were instructed to remember the digits in reversed order. Maximum length of the string was eight digits and again each correctly remembered string accounted for one point with a maximum of 14 and were included in the analyses as z-scores.

Reading comprehension. Reading Comprehension was measured using a standardized test for Grade 3 (Feenstra, Krom, & Van Berkel, 2007) and Grade 5 (Feenstra, 2009). Both tests consisted of two parts. The first part contained 25 multi-choice questions and the second part consisted of 30 multiple-choice questions. The second part was adapted to the reading level of each child measured in the first part; poor readers received an easier version than the good readers. The scores were transformed into respective age norms and thereafter transformed into z-scores, which enables across test and across grade comparisons. Normal average scores are 22 for Grade 3 and 45 for Grade 5 students.

Word decoding. Word decoding speed was measured using a word reading task (Verhoeven, 2005) that is administered twice a year at most Dutch primary schools. On the card 120 two- or three syllable words were presented, divided over four columns. Three versions are available, and version (B) used for the experiment was not recently administered at the schools. Children were instructed to read aloud as many words as possible within one minute. Every correctly read word was scored as a point and scores were included in the analyses as *z*-scores.

***Experimental texts*.** Three texts were constructed at different reading comprehension levels: Grade 1, Grade 3, and Grade 5. The texts were adapted from a standardized reading test to determine technical reading level (Jongen & Krom, 2009; Visser, Van Laarhoven, & Ter Beek, 1996). Minor adjustments were made to ensure that the length of the Grade 1 and 3 texts was equal (words n = 152). In order to match the length of the Grade 3 text to both the Grade 1 (n = 152) and Grade 5 (n = 232) text, two versions of the Grade 3 text were generated; a normal and an extended version.. This made sure students were involved in reading for about the same amount of time in order to control for concentration and motivational issues. In addition, one practice text at Grade 5 level was constructed and presented prior to the target texts.

To ensure an increase of difficulty from Grade 1 to Grade 5 texts, several text characteristics were considered. Measures of Lexical Richness (Vermeer, 2000) were calculated in order to determine the size of vocabulary needed for text comprehension. In addition, log transformed word frequency scores for every word was adapted from a Dutch child corpus (Tellings, Hulsbosch, Vermeer, & Van den Bosch, 2014) containing 11.5 million words and 5 million unique words from reading material (42% text books and tests, 38% books and magazine, and 20 % other media). Also number of words, number of sentences, mean sentence length, and mean syllable length were calculated. Table 1 shows an increase for all characteristics from Grade 1 to Grade 5 texts.

Table 1

Specific Characteristics for Target Texts A, B, and C

Text	Text A	Text B		Text C
Level	Grade 1	Grade 3		Grade 5
Version		Short	Extended	
Measure of Lexical Richness	3.23	4.03	3.62	4.83
Word frequency (log)	68.92	68.18	68.42	72.38
Number of function words	72	70	101	127
Number of content words	80	82	131	105
Mean word length in syllables	1.13	1.32	1.38	1.48
Number of sentences	25	20	31	20
Mean sentence length in words	6	7.5	7.4	11.5

Apparatus

The experiment was conducted using a Tobii T120 eye tracker with a sampling rate of 120 Hz. Participants were sitting in a chair adjusted to their height. The eye tracker was placed on a monitor arm at a distance of 70 cm. The eye tracker was set at the appropriate height in accordance with the head position of the child. A table with a button box was placed next to the participants.

Texts were presented on a 17 inch screen with a 1280 x 1024 resolution with a black background and white letters. Texts were presented 200 px from the sides of the screen in Arial 20 px roman style; a normal font type, which is not bold, underlined or cursive. The title was printed in bold. All sentences started at a new line.

Procedure

In the first phase of the study, students' reading comprehension, working memory and decoding speed were measured. The reading comprehension task was administered in class during two sessions. The first session lasted about 40 minutes and the second about 50 minutes. The working memory and decoding speed tasks were administered individually in one session of about ten minutes.

In the second phase, participants were positioned in front of the eye tracker, with their right hand on the two buttons. Participants were instructed to read the texts for comprehension and to recall the text afterwards to make sure the students concentrated on the task. Recall was free and children were asked what they remembered. The task consistently ended with the question 'is there anything else you remember?' Whenever the answer to this question was negative, the task stopped. All instructions were read aloud by the instructor and the children read along. After the instructions, the eyes were calibrated using nine red fixation dots on a black background. After reading and recalling the practice text, calibration was repeated before reading the first and before reading the second target text. The order of the texts was counterbalanced across participants. Phase two took approximately 30 minutes per participant.

Data analyses

Fixations were calculated with a minimum duration of 80 ms and a maximal dispersion of 1°. Areas of Interest (AOI) were determined by pixel positions of the words, taking into account an additional 5 px at the start of each new word. Finally, fixations with durations longer than 1200 ms were deleted, which was approximately 0.03% of the data.

Averaged reading times were calculated for each word (Hyöna, Lorch, & Rinck, 2003), including: a) *Gaze duration (G)*; the sum of fixation durations on the first encounter, b) *Look Back duration (LB)*; the sum of all fixations on previous text, c) *Second Pass duration (SP)*; the sum of fixation durations when reading the word for a second time (only possible when a regression was made). Furthermore, d) *Skipping probability (S)* and e) *Regression probability* were determined for each word by constructing a binomial variable that signified whether words were skipped or regressions were made or not. Mean probability scores represent the chance of a word being skipped or the chance regression to previous text segments occurs after fixation on a word. Measures of gaze, look back and second pass duration were log transformed.

To determine the role of student and text-related characteristics, we conducted mixed logit regression model for the probability measures and linear mixed effects regression models for the reading time measures (LMER). A backward stepwise selection procedure was used, deleting all predictors and interactions that did not reach significance at the level of 5% (Baayen, 2008). The full model contained main effects of text-related characteristics: grade (3 vs. 5), word class (Function vs. Content), and text difficulty (Easy vs. Difficult). Two-way interactions of text characteristics (Grade X Word Class, Grade X Text Difficulty, Word Class X Text Difficulty) and a three-way interaction of grade, word type, and text difficulty were entered into the model. Next, student-related characteristics were included using a forward stepwise selection procedure (Viebahn, Ernestus, & McQueen, 2012), comparing models with and without a particular skill. Predictors were included in the following order: decoding, short-term memory, working memory, and reading comprehension skill. Lastly, interactions of each student variable (decoding, short-term memory, working memory, and reading comprehension skill) and grade were tested.

For the single word analyses (skipping rate, gaze duration, and second pass durations), effects of word length and frequency were included in the model. Finally, forward model comparisons - of the fitted and reduced models - based on log-likelihood ratio tests were conducted to determine the maximum random slope effect structure by participant and word for each model. Thereafter, the fitted model was re-examined and insignificant fixed effects were deleted. For mixed linear-effect models and mixed logit models, respectively *t*-values and *z*-values are reported.

Results

Descriptives

Table 2 depicts the means and *SD*s of the raw scores of the student characteristics for each grade: decoding skill, short-term memory, working memory, and reading comprehension. Differences between grades were found for decoding, $t\ (41) = 6.35$, $p < .001$, $d = -1.81$, and reading comprehension, $t\ (47) = 4.26$, $p < .001$, $d = -1.22$, but not for short-term, $t\ (44) = 1.71$, $p = .094$, $d = -0.49$, and working memory, $t\ (47) = 1.82$, $p = .075$, $d = -0.52$. Variables showed no multicollinearity (all VIF's were below 1.41). Mean skipping rates, reading time durations of function and content words as a function of grade and text difficulty are presented in Table 3.

Table 2

Mean raw Scores and Standard Deviations of Student-related Characteristics among 3rd and 5th Grade Students

Student characteristics	*Grade 3* *n = 24* *M* (*SD*)	*Grade 5* *n = 20* *M* (*SD*)	*t*	*p*
1. Decoding skill	56.21 (16.38)	81.25 (11.47)	5.94	< .001
2. Short-term memory	6.96 (1.43)	7.50 (1.15)	1.39	.171
3. Working memory	4.21 (1.06)	4.55 (0.94)	1.13	.266
4. Reading comprehension	29.33 (13.15)	42.25 (11.00)	3.55	< .001

Table 3

Mean Skipping and Regression Probabilities in Percentages and Reading Time Durations in ms for Function and Content Words as a Function of Grade and Text Difficulty (N=*44)*

		Function words					Content words				
		S	G	R	LB	SP	S	G	R	LB	SP
Text difficulty	Grade	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Easy	Grade 3	42.3 (9.5)	344 (59)	16.6 (4.3)	947 (282)	560 (180)	39.6 (6.9)	365 (52)	15.8 (5.9)	694 (196)	495 (189)
	Grade 5	65.4 (14.3)	280 (33)	10.6 (5.3)	608 (267)	325 (96)	33.9 (19.3)	309 (48)	21.6 (10.0)	511 (184)	322 (86)
Difficult	Grade 3	52.3 (15.7)	340 (58)	9.3 (4.4)	611 (195)	366 (142)	24.8 (10.5)	396 (68)	20.8 (6.9)	628 (215)	413 (157)
	Grade 5	69.2 (4.9)	284 (33)	17.2 (7.1)	1782 (492)	343 (105)	46.2 (9.0)	360 (59)	18.0 (7.5)	739 (200)	346 (119)

Note. S = Skipping Probability. G = Gaze duration. R = Rereading Probability. LB = Lookback duration. SP = Second pass duration. Scores reported in this table are based on participant means.

2

Skipping probability

For skipping probability, a mixed logit regression model analysis was run on score for each single word in the text, which was either scored as being read or skipped. The total amount of trials was 15.280, and 43.45% of all words were skipped. Results on the fitted model are presented in Table 4.

Table 4

Results on the Statistical Analysis of Skipping Probability

Predictor: Fixed effects	*ß*	*z*	*p*
Intercept	1.495	1.737	= .082
Word length	0.619	6.406	< .001
Grade	-0.417	-2.232	= .002
Word class	-3.560	-3.976	< .001
Text difficulty	-0.921	-0.847	= .397
Grade: word class	0.700	4.984	< .001
Grade: text difficulty	0.053	0.304	= .761
Word type: text difficulty	6.224	4.217	< .001
Grade: word class: text difficulty	-0.937	-3.919	< .001
Decoding	-0.326	-2.526	= .012
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.172	572.53	< .001
Word	1.429	2409.8	< .001
Particpant: decoding	0.281	10.757	= .005
Word: decoding	0.184	46.288	< .001
Word: text difficulty	0.320	26.82	< .001

The analysis showed main effects for word length, indicating that shorter and more frequent words are more often skipped. Also, a main effect of grade was found, showing that Grade 5 students more often skip words. Further, the main effect of word class showed that function words are more often skipped than content words. No main effect for text difficulty was found. No interactions with grade and student characteristics were found, but the three-way interaction effect of grade, word class and text difficulty was significant

Further exploration of the three-way interaction using mixed logit effect models showed that Grade 3 students skip function words more often (M = 52.3%) than content words (M = 24.8%) in *difficult* texts, $ß$ = 1.166, z = 6.208, p < .001, but no differences were found for *easy* texts, $ß$ = 0.067, z = 0.242, p = .808 ($M_{function}$ = 42.9 %; $M_{content}$ = 39.6%). For Grade 5 students; fifth graders skip function words more often (M = 67.5%) than content words (M = 39.2%) in both *easy*, $ß$ = 1.299, z = 8.454, p < .001, and difficult texts, $ß$ = 0.951, z = 3.936, p < .001. An overview of average percentages of skipping probabilities is presented in Figure 1.

2

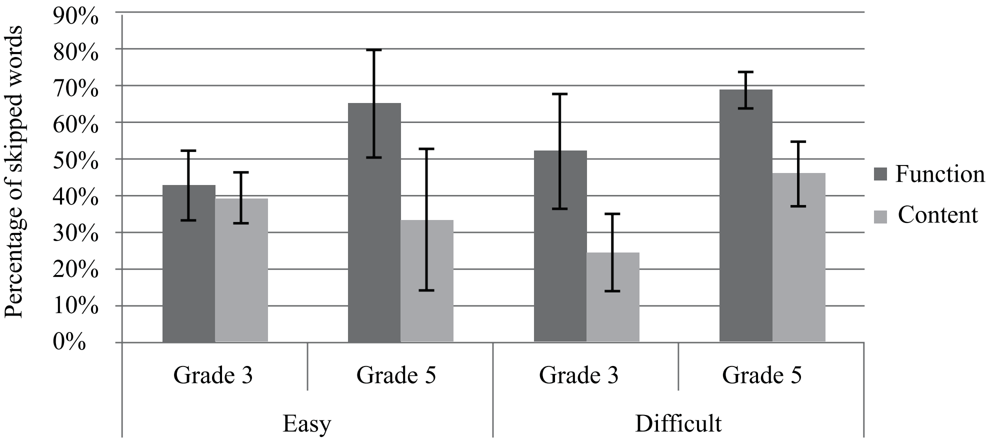


Figure 1. Interaction Effect of Grade, Word Class and Text Difficulty for Skipping Probability.

Gaze duration

To determine in what way text and student-related characteristics predict gaze durations, a mixed linear model was fitted on a dataset including all words that were read, resulting in total amount of 8024 trials. The results of the analysis on gaze duration are presented in Table 5.

Firstly, we found a main effect of word length; longer words show longer gaze durations. Also, the difficulty of the text influenced gaze duration and longer durations were found for more difficult texts.

With respect to student characteristics, main effects were found for decoding, short-term memory and working memory. Better decoding skills result in shorter gaze durations. On the one hand higher short-term memory resulted in shorter gaze durations, whereas higher working memory capacity results into longer gaze durations. No main effect of grade was found, but an interaction of working memory and grade was significant. Further exploration of the interaction showed that although both Grade 3 students' gaze durations were relying on working memory, $ß = 0.040$, $t = 1.99$, $p = .047$, but Grade 5 students were, $ß = -0.060$, $t = -3.01$, $p = .026$, the direction of the effect was reversed. These effects indicate that 3rd graders gaze durations are slower when working memory capacity is higher and for 5th graders gaze durations are faster when working memory is higher.

Table 5

Results on the Statistical Analysis of Gaze Duration

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	5.568	47.67	< .001
Word length	0.014	7.61	< .001
Text difficulty	0.030	2.02	= .043
Grade	-0.013	-0.69	= .489
Decoding	-0.081	-4.30	< .001
Short-term memory	-0.044	-3.15	= .002
Working memory	0.260	3.11	= .002
Grade: working memory	-0.045	-3.19	= .001
Predictor: Random Effects	*Explained variance*	χ^2	*p*
Participant	0.006	880.24	< .001
Word	0.003	584.45	< .001

Regression probability and look Back duration

For regression probability, a mixed logit regression model analysis was run on all the words that were read in first pass. For all words, it was determined whether participants looked back after reading this word or not. The results are presented in Table 6. The results showed a main effect of grade, indicating that the chance that 3rd graders make regressions is smaller than for 5th graders. In addition, a two-way interaction of word class and text difficulty was found, indicating that although in both the *easy* ($ß = 0.656$, $z = 4.238$, $p < .001$, $M_{function} = 13.2$ %, $M_{content} = 19.2$ %) and *difficult* ($ß = 1.025$, $z = 5.888$, $p < .001$, $M_{function} = 10.3$ %, $M_{content} = 18.9$ %) texts regressions are more frequent for content words, this effect is larger for difficult texts.

2

Table 6

Results on the Statistical Analysis of Regression Probability

Predictor: Fixed effects	*ß*	*z*	*p*
Intercept	-2.557	-4.430	< .001
Grade	0.223	2.461	= .014
Word class	-0.076	-0.485	= .627
Text difficulty	-0.218	-1.546	= .122
Word class: text difficulty	0.317	1.980	= .048
Predictor: Random Effects	*Explained variance*	χ^2	*p*
Participant	0.244	236.37	< .001
Word	1.923	1030	< .001

For look back duration, a mixed linear effect model was fitted, including 2376 trials. Hence in 29.6% of the cases students were involved in looking back. A summary of the fitted model can be found in Table 7. A main effect of text difficulty was found, indicating that students spent less time looking back to previous text parts in *easy* text (*M* = 662 ms) compared to *difficult* texts (*M* = 972 ms).

Table 7

Results on the Statistical Analysis of Look Back Duration

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	5.988	144.42	< .001
Text difficulty	0.133	2.78	= .005
Predictor: Random Effects	*Explained variance*	χ^2	*p*
Participant	0.025	61.482	< .001
Word	0.192	920.31	< .001
Word: text difficulty	0.153	21.257	< .001

Second pass duration

A mixed linear effect models was also fitted for second pass duration. Total amount of trials was 1311, indicating that students, after regressing to previous text parts, reread the word from which they regressed in more than half (55.2%) of the cases. Results of the fitted model are presented in Table 8.

With respect to text-related characteristics, a negative main effect was found for text difficulty, indicating that more difficult text showed shorter second pass durations. Moreover, a two-way interaction of grade and text difficulty was significant. Mean look back durations for each grade and text are presented in Figure 2. The interaction showed that second pass durations for 3rd grade students were longer for *easy* texts than for *difficult* texts, $ß$ = -0.168, t = -2.42, p = .002, but this is not true for 5th graders, $ß$ = 0.043, t = -0.85, p = .394. No student-related characteristics significantly contributed to the fitted model.

Table 8

Results on the Statistical Analysis of Second Pass Duration

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	6.744	29.118	< .001
Grade	-0.165	-4.366	< .001
Text difficulty	-0.719	-2.663	= .008
Grade: Text difficulty	0.111	2.511	= .012
Predictor: Random Effects	*Explained variance*	χ^2	*p*
Participant	0.028	53.92	< .001
Word	0.063	140.19	< .001

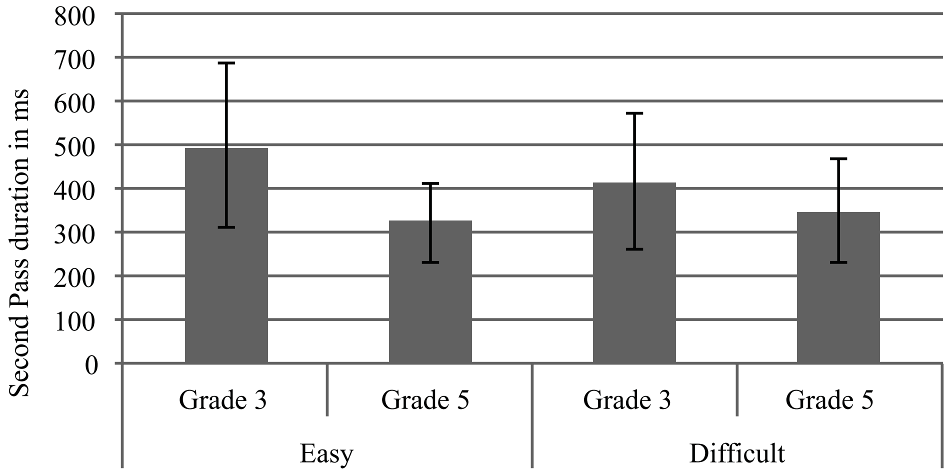


Figure 2. Interaction Effect of Grade and Text Difficulty for Second Pass Duration.

Discussion

In this study, we examined eye movements of Grade 3 and 5 students to determine the interaction of text and reader characteristics. The results showed that both text and reading characteristics influence eye movements. Text-related characteristics were found to interact with grade in skipping probability and second pass duration, but not for gaze duration and look back duration, showing that 3rd grade students skipped function words more often in difficult texts and 5th grade students in easy texts compared to content words. Also 3rd graders spend more time in second pass in easy texts compared to 5th grade students. Moreover, student-related characteristics are found to be particularly important for gaze duration. In addition, interactions of grade and working memory were found for gaze duration measures, indicating that working memory has a positive effect on reading times of 3rd graders, but a negative effect on reading times of 5th graders. These results suggest that skipping probability is relying mostly on text-based characteristics, gaze durations are also relying on students' characteristics and look back and second pass durations change throughout grades.

The results of this study lend support to our first hypothesis: text characteristics influence eye movement patterns. Firstly, length predicted skipping probability, showing that short words are skipped more often. Also, gaze durations were predicted by word length, which is in line with previous finding on adults (Rayner & Liversedge, 2011) and children (Hyönä & Olson, 1995; Joseph et al., 2009). Word frequency did not show such an effect, which is in line with Blythe et al. (2006), but not in line with many other studies showing frequency effects (Blythe et al., 2009; Joseph et al., 2013). One of the main reasons for not finding word frequency effects could be that word type and frequency are highly related. A t-test showed word frequency to be lower for content (M = 102.20, SD = 26.01) compared to function word (M = 64.71, SD = 46.25). However, post-hoc analyses showed the additional value of word type over and above word frequency and that the previously observed effect for word frequency disappeared. Moreover, as no effects of word length and frequency effects were found for second pass durations, this suggests that these effects disappear after initial processing.

Secondly, text-related characteristics were found to influence processing. First, function words and words in easy texts were indeed found to be skipped more often, confirming the results of Roy-Charland et al. (2007), although the present study does not rule out predictability effect. Second, no main effects of word class were found for reading time measures. This suggests that word class influenced skipping, but not reading time durations. Future studies could investigate whether processing is affected by the distribution of content and function words across texts, because text with relatively more function words (which was the case in the Grade 5 text) might halt skipping. Finally, text difficulty predicted - apart from skipping probability - also gaze and second pass duration, indicating that text integration seems to be more effortful in more difficult texts, resulting into longer reading times.

With respect to regressions, we confirmed our hypothesis that 5th graders more often look back, although regression rate are much larger than expected and reported by Rayner (1985). This is most likely caused by reading whole text with relatively few consecutive comprehension questions. The results of the present study do not show regression probability is indeed higher for older students, which indicates proficient readers are indeed better at monitoring their comprehension (Garner & Reis, 1981; McNamara & O'Reilly, 2009). However the duration of look backs is not different for 3rd and 5th graders, but was relying on the level of difficulty of the text. Monitoring behaviour is hence causing more, but not longer regressions. Older readers might hence be better at locating the information needed to solve their comprehension problem. The results of this study do not indicate that shorter look back reading times are related to short-term memory, which is conflicting with hypotheses on allocating information in text. In order to gain more insight in these phenomena, future research should focus on the exact time course of the regressions.

Partial support was also found for our second set of hypotheses which stated that student characteristics also affect eye movements. This study showed that the inclusion of reader characteristics improved the model for skipping probability and gaze duration. Decoding predicts skipping probabilities and gaze durations over and above other measures, including age (grade), which is in line with our prediction that lower level skills influence early reading processes. The absence of any effects of decoding on second pass might be either caused by a lack of statistical power, as the amount of trials is limited, or due to the fact that these processes are solely influenced by higher-order skills. These results are also in line with previous behavioural measures, showing that better decoding skills (Verhoeven &Van Leeuwe, 2008) and older students have faster reading times (McConkie et al, 1991).

Although we expected reading comprehension skills to contribute to predicting eye movements, we did not find an additional value for reading comprehension skills over and above grade and sub skills (decoding, short-term and working memory). This might indicate that the reading comprehension test that we used was most likely not tapping into higher level linguistics and cognitive skills. It might also be the case that the texts used in this study were too easy and hence did not enhance higher level processing. For this reason, future studies should not only focus on the differences between skilled and less skilled readers, as this study shows more fine grained effects on the subskills involved in reading comprehension are much more informative.

Of particular interest are the effects of memory in gaze duration. Both short-term memory and working memory were predicting the amount of time students needed to read a word on the initial encounter, which is contradictory to our hypothesis. The results can be explained in lines of mental model building, because memory of previous text is important in reading new information (Kintsch, 2004; Van de Broek, Rapp, & Kendeou, 2005). It can be assumed that linking information is only easy when the relevant information is available in memory. If this is not the case, the reader experiences a cognitive overload (Van Merriënboer & Sweller, 2005). In particular, differences in working memory are found among grades; although working memory is positively associated with gaze duration in 3rd grade, it is negatively related for larger for 5th graders. Concluding, these results suggest that working memory is important for comprehension processes, but that the direction of the effect changes throughout the grades.

Several limitations of this study should be addressed at this point. First, as this study aimed at exploring the role of text characteristics across grades in a natural reading environment, the sample frequency causes limitations with respect to the temporal resolution of the data. Especially with respect to look back and second pass duration, the limited amount of samples might cause a higher temporal sampling error (Andersson, Nyström, & Holmqvist, 2010) and might have influenced the reported effects. Following Andersson et al.'s (2010) calculations, we are confident that the temporal sampling error is reduced to a similar level as a 1000 Hz eye tracker, taking into account the large amount of data points that we have included in the analyses.

Furthermore, conclusions based on text difficulty should be considered with care. The texts used in this experiment were regarded as being relatively easy or difficult on the basis of specific text characteristics. However, this measure is not directly related to the perceived text difficulty or the coherence of the text; it is unclear if students did indeed experience the text as being easier or more difficult. Future studies should address this issue by adding attitude questions on the target texts, and further exploring what type of text characteristic causes the regressive behaviour.

To conclude, the results of this eye movement study showed that not only text characteristics and grade (age), but students characteristics should be considered when conducting eye-tracking studies. In addition, this study shows that as students progress to higher grades, they do not only skip more words, become faster readers and look back more often, working memory seems to play a more and more important role in reading comprehension processes. These results are valuable for instructional designs. Cognitive load of instructions should be reduced in order to optimise reading comprehension processes especially in higher grades.

2

References

Andersson, R., Nyström, M., & Holmqvist, K. (2010). Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. *Journal of Eye Movement Research*, *3*(3), 6, 1-12.Baayen, R.H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Benjamin, R. G. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. Educational Psychology Review, 24(1), 63-88. doi: 10.1007/s10648-011-9181-8.

Blythe, H. I., & Joseph, H. S.S.L. (2011).Children's eye movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye-movements*. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199539789.013.0036.

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. doi:http://dx.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. doi:10.1016/j.visres.2009.03.015.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye-movements in reading: Perceptual and language processes* (pp. 275-307). New York, NY: AcademicPress.

Chamberland, C., Saint-Aubin, J., & Légère, M. (2013). The impact of text repetition on content and function words during reading: Further evidence from eye movements. *Canadian Journal of Experimental Psychology*, *67*(2), 94-99. doi:10.1037/a0028288.

Clifton, C., Jr., & Staub, A. (2011). Syntactic influences on eye movements during reading. In S. Liversedge, Iain Gilchrist, and Stefan Everling (Eds), *Oxford Handbook on Eye-movements. Chapter 49* (pp. 895-909). Oxford: Oxford University Press.

De Abreu, P. M. J., Gathercole, S. E., & Martin, R. (2011). Disentangling the relationship between workingmemory and language: The roles of short term storage and cognitive control. *Learning and Individual Differences*, *21*(5), 569-574. doi:10.1016/j.lindif.2011.06.002.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension:A meta-analysis. *PsychonomicBulletin & Review*, *3*(4), 422-433. doi:10.3758/s13423-013-0565-2.

Feenstra, H. (2009). *Begrijpend lezen groep 7* [Reading comprehension Grade 5]. Arnhem: Cito.

Feenstra, H., Krom, R., & Van Berkel, S. (2007). *Begrijpend lezen groep 5* [Reading comprehension

Fromkin, V.A. (2000). *Linguistics: An introduction to linguistic theory*. Malden, MA: Blackwell Publishing Ltd.

Garner, R., & Reis, R. (1981). Monitoring and resolving comprehension obstacles: An investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders. *Reading Research Quarterly*, *16*(4), 569-582. doi:10.1080/17470211003592613.

Graesser, A., McNamara, D., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers.* 36*(2),* 193-202. *doi:* 10.3758/BF03195564.

Hautala, J., Hyönä, J., & Aro, M. (2011). Dissociating spatial and letter-based word length effects observed in readers' eye movement patterns. *Vision Research*, *51*(15), 1719–1727. doi: 10.1016/j.visres.2011.05.015.

Häikiö, T., Bertram, R., & Hyönä, J. (2010). Development of parafoveal processing within and across words in reading: Evidence from the boundary paradigm. *The Quarterly Journal of Experimental Psychology, 63(*10*),*1982-1998. doi:10.1080/17470211003592613.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology, 102,* 167-181. doi:10.1016/j.jecp.2008.04.002.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 417-429. doi:10.1037/0278-7393.16.3.417.

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. doi: http://dx.doi.org/10.1016/j.visres.2009.09.012.

Hyönä, J. (2011). Foveal and parafoveal processing during reading. In S. Liversedge, I. Gilchrist & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 819-838). Oxford: Oxford University Press.

Hyönä, J., Lorch, R.F., Jr., & Rinck, M. (2003). Eye-movement measures to study global text processing.In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313-334). Amsterdam: Elsevier Science.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: effects of word-length and word-frequency. *Journal of Experimental Psychology: Learning Memory and Cognition, 21,* 1430-1440. doi:10.1037/0278-7393.21.6.1430.

Jongen, I., & Krom, R. (2009). *DMT en AVI: AVI Handleiding* [DMT and AVI: AVI manual]. Arnhem: Cito.

Joseph, H., & Liversedge, S. (2013). Children's and adults' on-line processing of syntactically ambiguoussentences during reading. *PLoS ONE, 8*(1), e54141. doi:10.1371/journal.pone.0054141.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., Gathercole, S. E., & Rayner, K. (2008). Children's and adults' processing of anomaly and implausibility during reading: Evidence fromeye movements. *The Quarterly Journal of Experimental Psychology, 61,* 708-723. doi:10.1080/ 17470210701400657.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length andlanding position effects during reading in children and adults. *Vision Research, 49,* 2078-2086. doi:10.1016/j.visres.2009.05.015.

Joseph, H. S. S. L., Nation, K., & Liversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. S*chool Psychology Review, 42*(2), 207-222. doi:10.1371/journal.pone.0054141.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329-354. http://dx.doi.org/10.1037/0033-295X.87.4.329.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122-149. doi: 10.1037/0033-295X.99.1.122.

Kaakinen, J., Hyönä, & Keenan, J. M. (2002). Perspective effects on online text processing. *Discourse Processes*, *33*(2), 159-173. doi:10.1207/S15326950DP3302_03.

Kaakinen, J., Hyönä, 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*, 447-457. doi:10.1037/0278-7393.29.3.447.

Kintsch, W. (2004). The Construction-Integration Model of Text Comprehension and Its Implictions for Instruction. In: R.B. Ruddell, & N. J. Unrau, (Eds.), *Theoretical Models and Processes of Reading*. *Fifth Edition* (pp.1270-1328). Newark, USA: International Reading Association.

Kort, W., Schittekatte, M., Compaan, E. L., Bosman, M., Bleichrodt, N., Vermeir, G., Resing, W. C. M., & Verhaeghe, P. (2005). *WISC-III NL. Handleiding* [Manual]. *Nederlandse bewerking* [Dutch version]. London: The Psychological Corporation.

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. doi:10.1016/ j.jml.2011.03.002.

Lester, A. L., Nagle, R. J., Johnson, G., & Fisher, D. (1979). Eye-movements dynamics of good and poor readers then and now. *Journal of Reading Behavior*, *9*(4), 319-328. doi:10.1080/ 10862967909547338.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movement during reading. In J. F. Stein (Ed.), *Vision and visual dyslexia* (pp. 251-262). London: Macmillan Press.

McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*(1), 1-43. doi:10.1.1.20.3312.

McNamara, D.S., & O'Reilly, T. (2009). Theories of comprehension and comprehension skill: Knowledge and strategies versus capacity and suppression. In A.M. Columbus (Ed.), *Advances in Psychology Research Vol. 62*. Hauppauge, NY: Nova Science Publishers.

Molfese, V. J., Molfese, D. L., & Modgline, A. A. (2001). Newborn and Preschool Predictors of Second-Grade Reading Scores: An Evaluation of Categorical and Continuous Scores. *Journal of Learning Disability*, *34*(6), 545-554. doi: 10.1177/002221940103400607.

Nation, K. (2007). Children's reading comprehension difficulties. In M. J. Snowling, & C. Hume (Eds.), *The Science of Reading: a Handbook* (pp. 248-265). Oxford: Blackwell Publishing Ltd.

Navab, A., Gillespie-Lynch, K., Johnson, S., Sigman, M., & Hutman, T. (2012). Eye tracking as a measure of responsiveness to joint attention in infants at risk for autism. *Infancy, 17*(4), 416-431. doi: 10.1111/j.1532-7078.2011.00082.x.

Oakhill, J., & 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. doi: 10.1207/s15327574ijt0403_3.

Radach, R., & Kennedy, A. (2013). Eye-movements in reading: Some theoretical context. *The Quarterly Journal of Experimental Psychology*, *66*(3), 429-452. doi:10.1080/ 17470218. 2012. 750676.

Raven, J.C. (1960). *Guide to the standard progressive matrices: Sets A, B, C, D and E.* UK, London: H.K. Lewis & Co. Ltd.

Rayner, K. (1985). The role of eye movements in learning to read and reading disability. *Remedial and Special Education, 6*(6), 53-60. doi: 10.1177/074193258500600609.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*, 211-236.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye-movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241-255. doi:10.1207/s1532799xssr1003_3

Rayner, K., Juhasz, B. J., & Polatsek, A. (2005). Eye-movements during reading. In M. J. Snowling, & C. Hulme (Eds), *The science of reading: A handbook* (pp. 79-97). Oxford: Blackwell.

Rayner, K., & Liversedge, S.P. (2011). Linguistic and cognitive influences on eye movements during reading. In S.P. Liversedge, I.D. Gilchrist, & S.Everling (Eds.) *The Oxford handbook of eye movement* (pp. 751-766). United Kingdom, Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199539789.013.0041.

Rayner, K., & Slattery, T. J. (2009). Eye-movements and moment-to-moment comprehension processes in reading. In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond Decoding: The behavioral and biological foundations of reading comprehension* (pp. 27-45). The Guilford Press, New York, NY.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in rea-ding: Comparisons to other models. *Behavioral and Brain Sciences*, *26*, 445-526. doi:10.1017/ S0140525X03000104.

Reynolds, R.E. (2000). Attentional resource emancipation: Toward understanding the interaction of word identification and comprehension processes in reading. *Scientific Studies of Reading*, *4*, 169-195. doi:10.1207/S1532799XSSR0403_1.

Roy-Charland, A., Saint-Aubin, J., Klein, R. M., & Lawrence, M. (2007). Eye-movements as direct tests of the GO model for the missing-letter effect. *Perception & Psychophysics, 69*(3), 324-337. doi: 10.3758/BF03193753.

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), 877-896. doi: 10.1037/a002373.1

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360-407.

Swanson, H. Lee., & Ashbaker, M. H. (2000). Working memory, short term memory, speech rate, word recognition and reading comprehension in learning disabled readers: does the executive system have a role? *Intelligence*, *28*(1), 1-30. doi:10.1016/S0160-2896(99)00025-2.

Tellings, A., Hulsbosch, M., Vermeer, A. & Van den Bosch, A. (2014). BasiLex: an 11.5 million words corpus of Dutch texts written for children. *Computational Linguistics in the Netherlands Journal, 4, 191-208.*

Van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist approaches in accounts of reading comprehension. *Discourse Processes, 39,* 299-316. doi:10.1080/0163853X.2005.9651685.

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. doi:10.1007/s11145-011-9337-4.

Van der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. M. D. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal in Research in Reading*, *31*(2), 203-223.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2). doi:10.1007/s10648-005-3951-0.

Vauras, M., Hyönä, J., & Niemi, P. (1992). Comprehending coherent and incoherent texts: evidence from eye movement patterns and recall performance. *Journal of Research in Reading*, *15*(1), 39-54.

Viebahn, M. C., Ernestus, M., & McQueen, J. M. (2012). Co-occurrence of reduced word forms in natural speech. In *Proceedings of INTERSPEECH 2012: 13th Annual Conference of the International Speech Communication Association* (pp. 2019-2022). Portland, Oregon, USA : [s.n.]

Verhoeven, L. (2005). *Drie-Minuten-Toets (DMT)* [Three-Minute-Test (TMT)]. Arnhem: Cito.

Verhoeven, L., & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology, 22*, 407-423. doi:10.1002/acp.1414

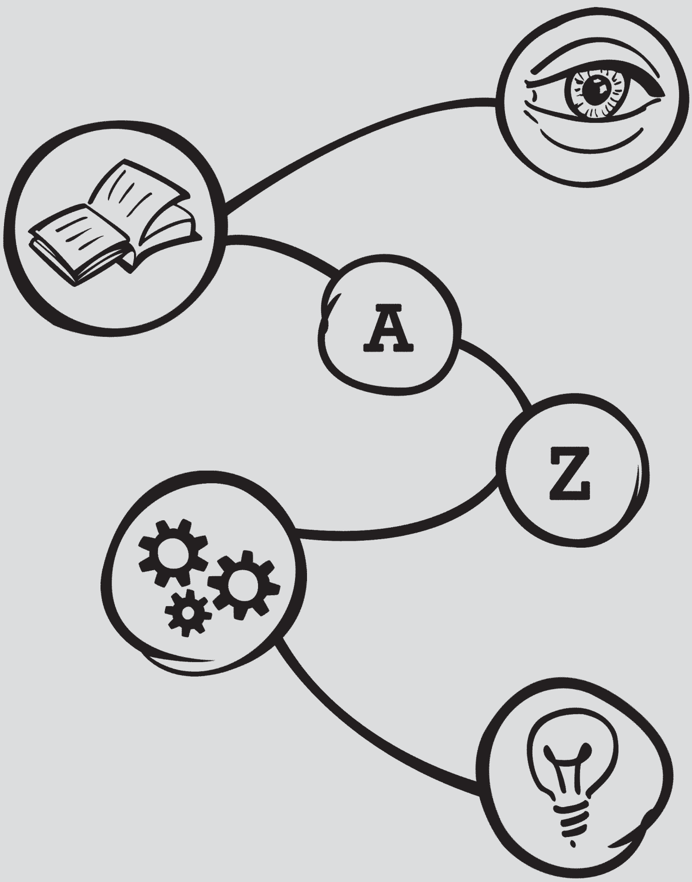
Vermeer, A. (2000). Lexicale rijkdom, tekstmoeilijkheid en woordenschatgrootte *[Lexical knowledge, text difficulty and vocabulary size]. Toegepaste TaalWetenschap in Artikelen*, *64*, 95-105.

Visser, J., Van Laarhoven, A., & Ter Beek, A. (1996). *Avi-toetspakket Handleiding* [AVI-testpackage Manual]. Den Bosch: KPC.

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. doi:10.1080/13506285.2014.881445.

CHAPTER 3

The effect of student-related and text-related characteristics on text comprehension: An eye movement study[2]



[2] This paper is based on:
De Leeuw, L., Segers, E., & Verhoeven, L. (resubmitted). The Effect of Student-Related and Text-Related Characteristics on Text Comprehension: An Eye Movement Study.

Abstract

The present study examined the role of eye movements in the relation between student characteristics (e.g., short-term memory, decoding efficiency, and vocabulary), text characteristics (wrap-up effects and text region) and text comprehension. Forty fourth graders read four explanatory texts and afterwards answered text comprehension questions. Skipping probability, gaze duration, regression probability, and regression path duration were examined. The results showed eye movements to moderate the effect of student characteristics on text comprehension. Eye movements may not only reflect efficient reading skills, but also compensational reading processes for lower-skilled students.

Introduction

In educational settings, new information is often acquired by reading expository text. In order to learn from text, readers need to build a mental model (Kintsch, 1994). The result is a text representation that can be stored in memory. Previous studies demonstrated that both the process and product of mental model building are related to children's abilities, such as word decoding (Huestegge Radach, Corbic, & Huestegge, 2009; Verhoeven & Perfetti, 2008), vocabulary knowledge (Calvo, Estevez, & Dowens, 2003), reading comprehension skills (Blythe & Joseph, 2011; Rayner, 1985; Reichle, Rayner, & Pollatsek, 2003), memory capacity (Daneman & Merikle, 1996; McNamara & O'Reilly, 2009; Swanson, Zheng & Jerman, 2009), and non-verbal intelligence (Tiu, Thomson, & Lewis, 2003). Also text-related characteristics have been found to influence reading comprehension processes, such as word length, word frequency (De Leeuw, Segers, & Verhoeven, 2015; Joseph, Nation, & Liversedge, 2013), wrap-up effects (Hirotani, Frazier, & Rayner, 2006; Rayner, Kambe, & Duffy, 2000) and text region (Hyönä, Lorch, & Kaakinen, 2002). However, it is still far from clear how reading processes influence the relation between student abilities and reading comprehension in children. Therefore, the present study examined the role of student-related and text-related characteristics, as well as eye movements (i.e. a reflection of the process of mental model building) on predicting reading comprehension (i.e. the product of mental model building).

A prerequisite for text comprehension is the construction of a coherent mental model (Kintsch, 2004). Coherent mental models are constructed during reading by means of constant updating of the current model (Kintsch & Van Dijk, 1978; Van der Broek Young, Tzeng, & Linderholm, 1999), which results in a 'network of propositions' (Kintsch, 1994: 295). Updating mental models is mainly done by creating links between the propositions with the help of inferences generated by either information within the current mental model (memory-based inferences) or prior knowledge (elaborate inferences) (Kintsch & Van Dijk, 1978; Van der Broek, Virtue, Everson, Tzeng, & Sung, 2002; Van den Broek, Rapp, & Kendeou, 2005). Research studying the time course of mental model building showed that readers usually update mental models at sentence boundaries (Blanc, Kendeou, Van den Broek, & Brouillet, 2008; Just & Carpenter, 1980), also called wrap-up effects.

Skills and capabilities of the reader also influence mental model building. More vocabulary knowledge helps the reader to better understand the concepts within the text, which in turn enhances the chance of memory-based inferences (Calvo et al., 2003;

Singer, Andrusiak, Reisdorf, & Black, 1992). As vocabulary is related to world knowledge, higher vocabulary also enlarges the chance of making elaborate inferences and linking the text to prior knowledge (Van den Broek, Lorch, Linderholm, & Gustafson, 2001). In a similar way, good readers make more inferences, because they are better at making inferences that span over larger text parts and because they make inferences using their background knowledge (McMaster, Espin, & Van den Broek, 2014).

As text comprehension is highly related to the number of inferences that are generated during reading (Van den Broek et al., 2001), short-term and working memory are also an important predictor of reading comprehension, both in adults (Daneman & Merikle, 1996) and children (Cain, Oakhill, Barnes, & Bryant, 2001; Cain, Oakhill, & Bryant, 2004). Especially developing readers' working memory can be (over)loaded with lower level processing (i.e. decoding and vocabulary) during text reading. And whenever readers are more involved in lower-level text processing, their working memory capacity available for higher-level text processing (Just & Carpenter, 1992), such as text integration, is reduced. Lastly, previous research has a relation of non-verbal intelligence and comprehension scores (e.g., Tiu et al., 2003).

Next to skills and capabilities of the reader, the reading process, reflected by eye movements, also relates to reading outcomes. A recent study with adult readers (Schotter, Tran, & Rayner, 2014) directly mapped eye movements onto comprehension outcomes. Schotter et al. (2014) showed that readers who made more regressions had a better understanding of the text. This could be interrelated with better monitoring behavior, as good readers are better at monitoring their comprehension during reading (McNamara & O'Reilly, 2009). However, compared to adult readers, developing readers are found to make more regressions (Rayner, 1990) and hence it remains unclear whether similar effects should be expected for these younger readers.

Finally, text-related effects may vary as a function of several student-related characteristics. First, word length and word frequency effects have been found to be smaller for adults compared to children (Joseph et al., 2013). Thus, developing readers have more difficulty with longer and infrequent words compared to adults. Second, wrap-up effects are different among individuals. High-school readers that were better at detecting inconsistencies, and therefore assumed to be better at reading comprehension, exhibited larger wrap-up effects than students that were not good at detecting inconsistencies (Schad, Nuthmann, & Engbert, 2012). These results indicates that integration at sentence final words is important for comprehension, as suggested by the Construction Integration model (Kintsch, 2004). Third, skilled developing readers spend more time on important text elements (Van der Schoot, Vasbinder, Horsley, & Van Lieshout, 2008). A similar effect was demonstrated in adults; readers that were better at writing summaries paid more attention to headings (Hyönä et al., 2002). Hence, readers vary in the way they allocate their attention to text segments.

Previous research has shown student characteristics to affect reading comprehension processes (Rayner, 1985) and products (Calvo et al., 2003; Singer et al., 1992). Although it is clear that process and product of mental model building are related (Kintsch, 2004; Van den Broek et al., 1999), it is still unclear how eye movements contribute to predicting text comprehension outcomes of developing readers. To understand in what way skills contribute to reading comprehension processes, we included several student-related characteristics (decoding skill, vocabulary knowledge, short-term memory, working memory, reading comprehension skill, and non-verbal intelligence). In addition, we included text-related characteristics (word length, word frequency, wrap-up and text region) and examined interrelations with readers' skills. Two research questions were addressed:

1. In what way are student-related and text-related characteristics associated with eye movements?
2. How are student-related characteristics, text-related characteristics, and eye movements associated with reading comprehension outcomes?

With respect to the first research question, it was hypothesized that student- and text-related characteristic both influence the reading process. We predicted large effects of word decoding efficiency on eye movements, because word decoding is highly related to the speed of reading, as are eye movement durations. With respect to wrap-up and text region effects, we expected readers with higher skills to spend more time on text integration (i.e., sentences final words), and more salient text regions (i.e., heading and first sentence of a paragraph). Further, working memory was expected to predict the occurrence of regression behavior and reading comprehension outcomes, because a small memory span limits the amount of information available for bridging inferences. With respect to the second question, no clear hypotheses was formulated, since very little research has focused on the effect of eye movements on reading comprehension outcomes in children.

Method

Participants

Students from two 4th grade classes from two Dutch primary schools participated. From the 48 participants, some (*n* = 5) were excluded due to unusable or missing fixation data or because the score on the text comprehension questions was more than two standard deviations from the mean (*n* = 3). In total, 40 students (13 girls, 27 boys, M_{age}= 9;4 years, age *range* 9;1-11;2) were included in the analyses. Participants had a normal IQ (*M* = 42.0, *SD* = 6.2, *Range* = 26-52) compared to a norm group of their age, all scoring above the 10th percentile (Raven, 2006).

Apparatus

The experiment was conducted using a Tobii T120 eye tracker with a sampling rate of 120 Hz. Spatial accuracy of this eye tracker is 0.5° and spatial resolution is 0.2°. For this reason, careful calibrations were obtained of which the quality was assessed by visual inspection. If the quality was poor, a recalibration procedure was started. All participants were sitting in a chair adjusted to their height. The eye tracker was placed on a monitor arm at a distance of 70 cm. The eye tracker was set at the proper height in accordance with the child's head position. A table with a button box and mouse was placed next to the participants.

Texts were presented on a 17 inch screen with a 1280 x 1024 resolution with a black background and white letters. Text-margins were 200 px from every sides of the screen. Font was Arial, 20 px and line height 3 in roman style. Headings were presented in a similar font, but the headings were printed in bold, with 30 px, line height 2, and subheadings in 20 px, line height 2.

Materials

Student-related characteristics

Decoding efficiency. Decoding efficiency was measured using a word reading task (Jongen & Krom, 2009) that is yearly administered at Dutch primary schools. On the card 120 words are presented, divided over four columns. Children were instructed to read aloud as many words as possible within one minute. Every correctly read word was scored as a point. The internal consistency of the test was rated as good (α = .94, Egberink, Janssen, & Vermeulen, 2009-2014)

Vocabulary knowledge. Vocabulary knowledge was tested by a standardized passive vocabulary knowledge test (*Leeswoordenschattaak*, Verhoeven & Vermeer, 1999).

This test consists of fifty multiple choice items in which each word was presented in a short and uninformative context, e.g., ‘He sells vegetables’. The students were asked for the meaning of the underlined word. Four multiple choice options were presented including a synonym of the target word, e.g., ‘grass’, ‘green soup’, ‘salad’, and ‘edible plants’. Two practice words were discussed prior to the test. Questions regarding the task were answered, though no hints to answers were given. Reported scores are the total number of correct answers with a maximum of 50. Reliability of the test is considerably good (α. = .87; Verhoeven & Vermeer, 1996).

Memory. A forward digit span memory task (*WISC-III^{NL}*, Kort et al., 2005) was administered in which the researcher read aloud a string of digits using a falling intonation and pausing one second between the digits. The students were instructed to remember the digits in same order as presented. The strings started short (n = 2) with two attempts for each string length. Whenever children correctly remembered at least one of two strings, the researcher continued with a longer string, adding one digit until a maximum (n = 9) was reached. Each correctly remembered string accounted for one point with a maximum of 16.

3

Second, a backward digit span memory task (*WISC-III^{NL}*, Kort et al., 2005) was administred. The researcher read aloud a string of digits using a falling intonation and pausing one second between the digits. The students were instructed to remember the digits in reversed order. The strings started short (n = 2) with two attempts for each string length. Whenever students correctly remembered at least one of two strings, the researcher continued with a longer string, adding one digit until a maximum (n = 8) was reached. Each correctly remembered string accounted for one point with a maximum of 14.

Third, a word span task (Verhoeven, Keuning, Horsels, & Van Boxtel, 2013) was administered. In this task, the researcher read aloud a string of high frequent CVC words with an one-second pause in between starting with two digits. Two strings of each length -using different words- were presented and thereafter the string was extended with a single word with a maximum of 7. Whenever the child repeated the string incorrectly four times in a row, the test was terminated. Each correctly recalled string accounted for 1 point, with a maximum of 12.

Finally, a *sentence repetition* task was administered, which measures the memory of syntactical information (Verhoeven, et al., 2013). The task consisted of Dutch sentences increasing in length and syntactic complexity. The research read aloud one sentence at a time and was instructed to repeat this sentence, paying attention to the words and their order. In total, the test consisted of 12 sentences. An error-free answer accounted for 2 points, when one mistake was made the item accounted for 1 point, and 2 or more mistakes resulted into no points. As soon as a student did not receive any points for four consecutive sentences, the test was terminated. Reported scores are the number of points on this test, with a maximum of 24 points.

A principal component analysis with varimax rotation was run on all memory measures. To determine the number of factors, a parallel analysis was run (O'Connor, 2000). Two factors were found (Eigenvalues: 1.365 and 1.085). The first factor showed high loadings on digit span forward (.706), word span (.841), and sentence span (.764), but not on digit span backwards (.050). The second factor showed a high loading on digit span backwards (0.994), but not on digit span forward (.085), word span (-.072), and sentence span (.093). Given these results, it can be concluded that the memory measures load on two factors; short-term memory (storage of information) and working memory (storage and manipulation of information). The loadings were used to calculate a weighted factor score for short-term memory bases and were included in the analysis.

Comprehension skills. Comprehension skills were measured using a standardized test for Grade 4 (Feenstra, 2008). This test consisted of two parts. The first part contained five text and 25 multiple-choice questions and the second part consisted of six text and 30 multiple-choice questions. Texts were both narrative and explanatory texts. A mixture of text-based and inference-based questions were included. Item response theory models were constructed based on a calibration experiment assessing the difficulty of each item. This enabled adaptive testing; the second part of the test was adapted to the reading level measured in the first part. Therefore poor readers received an easier version and the good readers received a more difficult version. Item response theory models were used to transform the results of the two tests into one scores that is related to the students' respective age (months of formal reading instruction), which enables across test comparisons. Reliability was good; for the easy version $\alpha = .84$ and for the difficult version $\alpha = .85$ (Egberink et al., 2009-2014).

Non-verbal intelligence. To assess non-verbal intelligence of the students, the *Standard Progressive Matrices* (Raven, 1960) test was administered. This multiple-choice test consists of 60 items which increase in difficulty. For each item, the student is asked to identify the missing element that completes the pattern shown in a specific figure. Items are divided over five sets (A, B, C, D, and E) with 12 items each. In set A and B, six answer options are presented, and in the other sets eight answers are provided. Prior to testing, the first and second items were discussed as an example. Every item was scored as a point with thus the maximum score was 60.

Experimental materials

Texts. Four texts were adapted from NieuwsbegripXL (CED-Groep, 2011), which is Dutch reading comprehension course that provides newspaper articles for children on a weekly basis. Topics of the target texts in this study were obesity, child labor, animal testing, and souvenirs. Each text consisted of five paragraphs, each presented on a separate screen. All paragraphs were preceded by a heading, which is standard for texts in this reading course and provided more power for analyzing headings.

Minor adjustments were made to ensure paragraph length was approximately similar. A summary of the characteristics for the reading material can be found in Table 1. In addition, one practice text was constructed and presented prior to the target texts. For each text, six subsequent multiple-choice questions on text likeability were administered, in order to clear the students' working memory. An example is 'How did you like this text?'. For this item, students answered on a five-point Likert-scale ranging from 1 (*not boring at all*) to 5 (*very boring*), though the labels of the scale differed for each question.

Table 1

Specific Characteristics for Target Texts

Text	Obesity	Child labor	Animal testing	Souvenirs
Word frequency (log)	239	250	250	244
Mean word length in characters	4.92	4.85	5.24	4.91
Number of sentences	48	46	48	48
Mean sentence length in words	7.77	8.33	7.85	7.79

Text comprehension. To test text comprehension, six multiple-choice comprehension questions were constructed for each text. Four of these questions could be literately deduced from the text (e.g. "In which area do we find child labor most frequently?" with answer options: a) agriculture, b) industry, c) stores, d) healthcare). The other two questions required the generation of an inference using two or more sentences. An example is: "Why do 60 children die each day? a) they do not have enough money to eat, b) they are being abused, c) they do not go to school, d) they breathe in dangerous dust." Reliability analysis showed one of the twenty-four questions to be unreliable and was therefore deleted from further analysis. The overall reliability of the remaining comprehension questions was good (α = .799).

Procedure

In the first phase of the study, students' skills were measured; decoding ability, vocabulary, short-term memory (digit/word/sentence), working memory, reading comprehension, and non-verbal intelligence. The vocabulary and comprehension task were administered in class within different sessions. The vocabulary test was administered in one session of about 15 minutes. The reading comprehension test was administered in two sessions. The first session lasted about 40 minutes and the second about 50 minutes. The non-verbal intelligence test lasted between 30-45 minutes. The decoding speed, short-term memory (digit/word/sentence) and working memory tasks were administered individually in one session of about 20 minutes.

The second phase was one eye tracking session. In a separate and quiet room, participants were positioned in front of the eye tracker, with their dominant hand on a button box. Participants were instructed to silently read the texts and answer questions afterwards. All instructions were read aloud by the instructor and the children read along. After instruction, the eyes were calibrated using nine red fixation dots on a black background. In order to get acquainted with the setup and navigation, an example text consisting of two pages was presented. Children were informed that they could navigate back and forth, though we must note that very few students actually navigated back. After reading, six likeability questions and two example multiple-choice text comprehension questions were presented on the screen. Each question was presented on a separate screen, and students were not allowed to navigate back to the text or to previous questions. After the instruction, the four target texts were read, starting with the calibration procedure prior to each text. After reading, the students were given six likeability questions and six multi-choice text comprehension questions. The order of the texts was counterbalanced across participants. The entire eye track session approximately 45 minutes per participant.

Results

Data analyses

Fixations were calculated with a minimum duration of 80 ms and a maximal dispersion of 1°. In order to analyze the eye movement data, every word within the text was considered as area of interest (AOI). Several characteristics of the AOI were included in the analysis, such as length (z-scores of the number of characters), word frequency (log transformed), the position in the sentence (dichotomous; 0 = non-final, 1 = final) and text segment (0 = remainder of the text, 1 = heading, 2 = first, 3 = final). Word frequency scores for every word was adapted from a Dutch child corpus (Tellings, Hulsbosch, Vermeer, & Van den Bosch, 2014) containing 11.5 million words and 5 million unique words from reading material (42% text books and tests, 38% books and magazine, and 20 % other media).

Fixations were deleted that were associated with moving the eyes to the beginning of the text and whenever they were longer than 1200 ms (0.72% of the data). Thereafter, four eye movement measures were calculated for each AOI (Juhasz & Pollatsek, 2011; Rayner, 1998): a) *Skipping probability* (S%); the chance a reader skips a words (binomial: 0 = read, 1 = skipped), b) *Gaze duration* (G); the sum of fixation durations in ms on the first encounter, c) *Regression probability* (R%); the chance a reader regresses from the target word (binomial: 0 = no regression, 1 = regression), *Regression* (R); the sum of all fixations in ms rereading previous text (including rereading of the target region), before progressing to the next word. No third or fourth passes were considered. All durational measures were log transformed and scores were deleted that deviated 2,5 SDs from the mean.

Separate models were run for skipping probability, gaze duration, regression probability and regression path duration. For all analyses, mixed effects regression models (*LMER*; Baayen, 2008) were run using the following procedure. First, a full model was created including all main fixed and random effects, as well as interactions among the fixed variables. A backward stepwise selection procedure was used[1], deleting all interaction effects that did not reach significance at the level of 5% on the ANOVA Wald test (car-package). In a next step, all non-significant main effects were deleted. Finally, random slope effects were added for the fixed effects (main and interaction) in the model, to account for intra-individual, -word, and -textual effects. Random slope structures were calculated by comparing unreduced and reduced models, based on log-likelihood ratio tests. The fitted model was re-examined and insignificant fixed effects were deleted if necessary. *Z*-values are reported for all logit linear models and *t*-values are reported for mixed linear-effect models.

[1] Similar results are found in a forward elimination procedure in which the reduced and full model were compared based on log-likelihood ratio tests.

Table 2

Correlations, Mean Scores, Standard Deviations and Range of Student-related Characteristics fourth Grade Students (N =*40)*

Student characteristics	1	2	3	4	5	6	7	*M*	*SD*	*Range*
1. Decoding skill	-							73.9	12.6	53-116
2. Vocabulary	-.020	-						20.0	3.6	10-27
3. Short-term memory	.002	.386*	-					12.5	2.00	7.8-16.1
4. Working memory	-.068	-.134	.193	-			4.10	1.2	2-7	
5. Reading comprehension	.094	.369*	.265*	-.071	-		32.3	12.7	6-60	
6. Non-verbal intelligence	-.256	.507**	.446*	.169	.450**	-	42.4	5.7	31-52	
7. Text comprehension	-.242	.221	.471*	.200	.523**	.258	-	18.7	3.4	8-23

Note. For text comprehension a combined score is reported here. For analyses, scores on the different texts were included. * $p < .001$ ** $p < .05$.

Descriptives

Table 2 depicts the correlations, means, SDs, and range of the student characteristics and text comprehension. Although some variables were moderately correlated, all VIF's were below 1.482, which indicates no problems with multicollinearity. Mean skipping probability, gaze duration, regression probability, and regression path duration for each region are presented in Table 3.

Table 3

*Means and Stand Deviations of Reading Time Durations in ms and Skipping and Regression probability in Percentages per Region (*N=*40).*

		S%	G	R%	R
Region	Wrap-up	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Heading	Non-final	47.3 (17.1)	352 (74)	25.2 (10.4)	1059 (257)
	Final	14.4 (15.9)	457 (136)	31.3 (13.0)	1127 (376)
First sentence	Non-final	43.4 (12.5)	360 (61)	35.4 (25.2)	1106 (516)
	Final	23.0 (14.9)	362 (75)	18.6 (14.0)	1016 (454)
Rest	Non-final	45.8 (10.0)	353(57)	21.1 (9.4)	979 (286)
	Final	31.4 (12.6)	355 (63)	28.6 (18.9)	1003 (421)
Final sentences	Non-final	51.1 (11.4)	336 (67)	29.9 (12.4)	1206 (575)
	Final	27.7 (17.1)	416 (99)	62.8 (25.4)	1555 (632)

Note. S%= Skipping Probability. G = Gaze duration. R% = Regression probability.
R = Regression path duration. Means are calculated on aggregated means per participant.

Effects of student-related and text-related characteristics on real-time reading behavior

Skipping probability.

To determine the effect of student-related and text-related characteristics on eye movements, a loglinear regression model analysis was run on the full dataset, including 42790 trials. The full model included random effects of participant, word, and text. Also, main fixed effects of student-related characteristics (word decoding, vocabulary knowledge, short-term memory, working memory, comprehension skill, and non-verbal intelligence) and text-related characteristics (length, frequency, wrap-up, text region) as well as two-way interactions of student-related and text-related characteristics were included.

Table 4

Results on the Statistical Analysis of Skipping Probability

Predictor: Fixed effects	*ß*	*z*	*p*
Intercept	-0.832	-9.890	< .001
Text-related characteristics			
Frequency	0.003	24.177	< .001
Heading	-0.649	-5.365	< .001
First sentence	0.041	0.571	= .568
Final sentence	0.083	1.579	= .114
Wrap-up	-.778	-13.734	< .001
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.185	1614.5	< .001
Participant: region[1]	0.081	57.074	< .001
Participant: position	0.0706	23.238	< .001
Word	0.322	1721.5	< .001
Text	0.002	428.85	< .001

Note. [1] averaged beta

The results of the final model are presented in Table 4. First, a positive main effect was found for word frequency, indicating that higher frequent words were skipped more often. Furthermore, a main effect was found for region, Wald test: $\chi^2 = 40.544$, $df = 3$, $p < .001$. Exploration of this main effect showed that words within the headings were skipped less often (30.8%) than in the remainder of the paragraph (38.6%), whereas words in the first (33.2%) and final (39.4%) sentences were not different. Finally, a main effect was found for wrap-up effects, indicating that that words at the final position were skipped less often (24.2%) than sentence non-final words (46.9%).

Gaze duration.

To determine the effect of student-characteristics on eye movements, a mixed linear regression model analysis was run on the gaze duration of each word in the text. About 56.5% of all words were read, resulting in a dataset of 24201 trials. The full model was identical to the one described for the skipping rate. Results of the fitted model are presented in Table 5. Random main effects were found for participant, word, and text. Also, a random slope was found for decoding efficiency within words.

Main fixed effects were found for student-related characteristics decoding and vocabulary, indication that higher decoding and vocabulary skills are related to shorter gaze durations. Furthermore, main effects of text-characteristics length, frequency, and region (Wald test: $\chi^2 = 29.595$, $df = 3$, $p < .001$) were found. The effects showed that longer words have longer gaze durations, whereas more frequent words have shorter gaze durations. With respect to text region, the results showed that student spent additional time in gaze duration on reading the heading and final clause of the paragraph compared to the remainder of the paragraph. However, the first clause did not show significant differences compared to the remainder of the paragraph.

Finally, an interaction of comprehension skill and region was found (Wald test: $\chi^2 = 10.007$, $df = 3$, $p = .019$). Further exploration of this interaction, with a median split on reading comprehension skill, showed an interaction for both skilled, $\chi^2 = 18.937$, $df = 3$, $p < .001$, and less skilled readers, $\chi^2 = 13.326$, $df = 3$, $p = .004$. This interaction showed that both skilled, $\beta = 0.139$, $t = 3.97$, $p = < .001$, and less skilled readers, $\beta = 0.083$, $t = 2.54$, $p = < .001$, spent additional time on reading the heading compared to the remainder of the paragraph and that the effect size of this interaction did not differ, $\beta = 0.056$, $t = 1.49$, $p = .136$. However, less skilled readers spent fewer time on reading the final sentence, $\beta = -0.037$, $t = -2.17$, $p = .031$, whereas this effect failed to reach significance for the skilled readers.

Table 5

Results on the Statistical Analysis of Gaze duration

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	5.703	130.14	< .001
Student-related characteristics			
Decoding	-0.083	-4.39	< .001
Vocabulary	-0.048	-2.21	= .022
Comprehension skill	-0.015	-0.74	= .637
Text-related characteristics			
Length	0.056	4.27	< .001
Frequency	-0.001	-17.49	< .001
Heading	0.110	4.26	< .001
First sentence	-0.009	-0.64	= .455
Final sentence	-0.031	-2.31	= .018
Interactions			
Comprehension skill: heading	0.050	2.60	= .009
Comprehension skill: first sentence	0.019	1.98	= .047
Comprehension skill: final sentence	0.008	0.83	= .405
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.0126	902.58	< .001
Word	0.0121	596.11	< .001
Word: decoding	0.0001	6.677	= .035
Text	0.0007	34.9	< .001

Regression probability.

To determine the effect of student-characteristics and text-related characteristics on regression probability, a logit mixed regression model analysis was run on all words that were read, resulting in 24201 trials. The full model included random effects of participant, word, and text. Also, main fixed effects of student-related characteristics (word decoding, vocabulary knowledge, short-term memory, working memory, comprehension skill, and non-verbal intelligence), text-related characteristics (wrap-up, text region) as well as two-way interactions of student-related and text-related characteristics were included. Note that word length and word frequency effects were not examined, as regions that are related to regressive eye movements (looking back to previous text segments) can vary in length and frequency.

The results of the final model are presented in Table 6 and show random effects for participant, word, and text. A negative main student-related fixed effects was found for decoding, indicating that higher decoding efficiency were related to less regressions. With respect to text-related characteristics, a main effect was found for wrap-up; regressions were more often initiated for sentence final words. Finally, two interaction effects were found: both decoding and non-verbal reasoning were found to be related to wrap-up effects, showing that wrap-up effects were larger for children with higher decoding and non-verbal intelligence.

3

Table 6

Results on the Statistical Analysis of Regression Probability

Predictor: Fixed effects	*ß*	*z*	*p*
Intercept	-1.162	-12.870	< .001
Student-related characteristics			
Decoding	-0.274	-3.258	< .001
Non-verbal intelligence	-0.034	0.409	= .683
Text-related characteristics			
Wrap-up	0.351	7.731	< .001
Interactions			
Wrap-up: decoding	0.120	2.623	= .009
Wrap-up: non-verbal intelligence	0.107	2.526	= .012
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.230	907.41	< .001
Word	0.211	396.23	< .001
Text	0.006	12.587	< .001

Regression path duration.

A mixed linear regression model analysis was run on the reading time for all words that triggered look back behavior (15.5% of the read words). In total, 3746 trials were included in the analysis. The full model was identical to the one of regression probability. Results of the final model are presented in Table 7. Random main effects were found for participant and word. Furthermore, main effects for decoding, wrap-up and text region were found. The main effect of decoding indicated that look back times were faster when decoding efficiency were higher. The main effect of region (Wald test: χ^2 = 32.524, *df* = 3, *p* < .001) showed longer reading times for the final region compared to the remainder of the paragraph. The heading and first sentence did not show significant effects.

In addition, two interaction effects were found in relation to decoding. First, decoding was found to interact with wrap-up. The interaction showed that higher decoding efficiency was related to shorter regression path durations on final regions. Finally, the interaction of the region and decoding (Wald test: χ^2 = 23.854, *df* = 3, *p* < .001) showed that students with low decoding efficiency spent more time looking back to previous regions than students with high decoding efficiency, but this is only true for the final region.

Table 7

Results on the Statistical Analysis of Regression Path Duration

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	6.777	265.71	<.001
Student-related characteristics			
Decoding	-0.14	-5.61	< .001
Text-related characteristics			
Wrap-up	0.044	1.75	= .032
Heading	0.008	0.12	= .986
First sentence	-0.030	-0.92	= .156
Final sentence	0.187	5.90	< .001
Interactions			
Wrap-up: decoding	-0.059	-2.20	= .028
Heading: decoding	0.035	0.56	= .578
First sentence: decoding	0.061	1.80	= .072
Final sentence: decoding	0.143	4.74	< .001
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.015	97.194	< .001
Word	0.022	69.651	< .001
Text: skipping	0.005	82.787	< .001

Effects of student-related and text-related characteristics on text comprehension

The second research question involved the relation of students' ability and their eye movements on text comprehension. A mixed linear effects regression model was run on a dataset of 42790 trials, including random effects of participant, word, and text. Further, main fixed effects of student-related characteristics (decoding, vocabulary, short-term memory, working memory, reading comprehension, and non-verbal intelligence), text-related characteristics (word length, word frequency, wrap-up, and text region), and eye movement measures (skipping probability, gaze duration, regression probability, and regression duration) were considered, as well as interactions among these variables.

The final model is presented in Table 8. Random effects were found for participant and text, as well as a random slope effect for text and skipping probability. A fixed main effect was found for short-term memory indicating that students with higher STM skills had higher scores on the comprehension questions. Furthermore, several interactions were found; an interaction of decoding and skipping probability and skipping probability and non-verbal intelligence and skipping probability.

Further exploration of the decoding interaction showed that skipping words was negatively affecting the results of the students for students that were in the lower comprehension group. However, for the group of students that scored relatively good on the comprehension scores, this effect seems to disappear; more skipping does not necessarily lead to worst scores on comprehension questions (see Figure 1a). For non-verbal intelligence, the interaction was found to be similar; skipping words had larger effects for low skilled students compared to their more skilled peers (see Figure 1b).

Table 8

Results on the Statistical Analysis of the Effect of Student-Related, Text-Related Characteristics and Eye Movements on Text Comprehension

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	- 0.037	-0.252	
Student-related characteristics			
Decoding	0.084	-0.592	= .682
Short-term memory	0.359	2.496	= .013
Non-verbal intelligence	0.091	-0.602	= .600
Eye movements			
Skipping	0.027	0.729	= .516
Interactions			
Decoding: skipping	0.052	7.259	< .001
Non-verbal intelligence: skipping	0.026	3.711	< .001
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.684	29412	< .001
Text	0.018	1173.4	< .001
Text: skipping	0.005	82.787	< .001

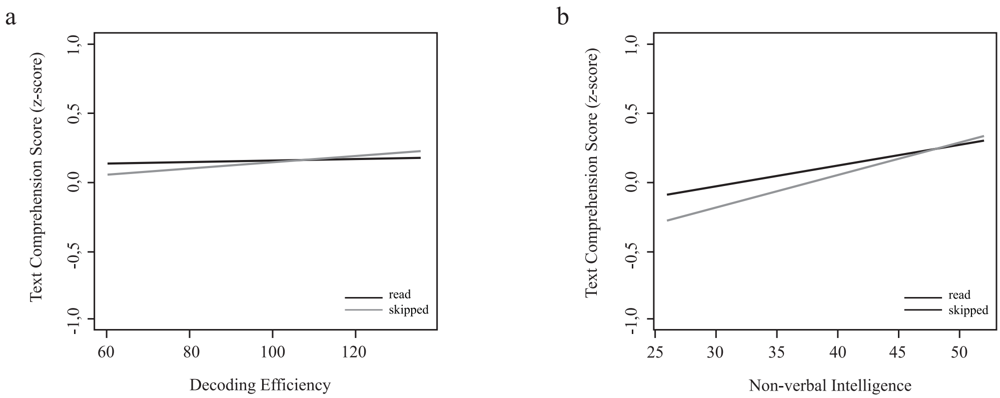


Figure 1. Interaction effect of decoding efficiency (1a) and non-verbal intelligence (1b) on text comprehension as a function of skipping probability

Discussion

The aim of this study was to determine the role of student-related and text-related characteristics on real-time processes on the one hand, and their association with text comprehension on the other hand. Regarding processing effects, this study showed decoding, vocabulary knowledge, and text-related characteristics to be related to eye movement measures. The effects on text comprehension showed that skipping probability moderates the effect of skills on comprehension and that text-related characteristics were not important in this respect.

With respect to the first research question, predictions regarding student-related effects on eye movement outcomes involved large effects of word decoding efficiency on eye movement measures, especially in early reading. Other literacy skills were expected to be of lesser importance. Interactions with text structure were expected, as experienced readers are more involved in strategic reading behavior (McNamara & O'Reilly, 2009). The results indeed showed strong effects for decoding efficiency on both gaze, regression path duration, and regression probability. Vocabulary was found to be related to gaze durations, but not to other eye movement measures. Other student-related skills were not found to be related to eye movements. These results are in line with previous studies showing faster reading times for skilled readers compared to less skilled readers (Blythe & Joseph, 2011; McMaster et al., 2014).

Furthermore, we hypothesized working memory to be related to regression measures, since a small memory span limits the amount of information available for bridging inferences and hence regressions are expected to be longer (Cain et al, 2001; 2004; Van den Broek et al, 2001). Nevertheless, we did not find evidence that regressions are depending on working memory (Swanson et al., 2009). This is in line with recent research on text reading, in which working memory effects for regressions were also absent in younger readers (De Leeuw et al., 2015). Nevertheless, we did find effects of short-term memory in gaze and skipping duration, indicating that memory is related to reading processes, but within earlier stages.

Further, we found several text-related characteristics to influence real-time processing. First, word length and word frequency effects for gaze duration were evidenced, which is in line with research showing longer and less frequent words to have longer reading times (Joseph et al., 2013). Second, clause wrap-up effects were found for regression measures, but not in gaze duration or skipping probability. This is partly in line with the literature, as Kaakinen and Hyönä, (2007) did find wrap-up effects, but only in gaze duration. Third, effects of text region were found for gaze duration and regression path duration. For paragraph headings and final sentences, longer gaze durations were found compared to the remainder of the text. The heading effect is similar to effects found for adults (Hyönä et al., 2002) and relates to effect of salience in children (Van der Schoot et al., 2008), although we fail to replicate an interaction with reading comprehension skill. The longer gaze durations in the final region is not in line with the study of Hyönä et al. (2002). These longer reading times could be contributed to reader fatigue (Graesser, Singer, & Trabasso, 1994; Schad et al., 2012; Van den Broek, Risden, & Husebye-Hartman, 1995) or mindless reading (Reichle, Reineberg, & Schooler, 2010), which is expected to influence developing readers to a larger extent than adults.

With respect to the second research question, effects of student-related, text-related and eye movements on text comprehension scores, as well as interrelations among these variables were explored. Memory capacity was expected to predict text comprehension, since it has been demonstrated that memory for text is facilitated by inference generation (Cain et al, 2001; 2004; Van den Broek et al, 2001). Our study indeed shows memory to be important, although we only found effects for short-term memory and not working memory. One explanation for this result might be our memory measures. For short-term memory we measured both verbal and non-verbal components of memory, whereas for working memory only one non-verbal component was included. Hence, it could be the case that the verbal component within the short-term measure loads high on comprehension and the lack of such a component in working limits its predictive value.

Furthermore, several interactions of skipping probability with skills (decoding and non-verbal intelligence) were found. Two conclusions could be drawn. First, children's reading comprehension processes are different from those of adults, as this study does not confirm the association of regression path durations and reading comprehension found in adults (Schotter et al., 2015). It seems that younger readers' comprehension is mainly regulated by initial processing, and not by monitoring behavior reflected by regressions. Second, eye movements (skipping probability) moderate the effect of the students' ability on reading comprehension. The results suggest that some less skilled readers adjust their reading (i.e., spend more time) to resulting in higher comprehension scores, whereas less skilled readers that fail to compensate for their lack of skill will obtain lower comprehension scores.

Several limitations of this study should be addressed at this point. First, as this study aimed natural reading environment, the temporal and spatial resolution of the data is limited. Following Andersson, Nyström, and Holmqvist's (2010) calculations, we are confident that the temporal sampling error is reduced to a similar level as a 1000 Hz eye tracker, taking into account the large amount of data points that we have included in the analyses. Nevertheless, problems related to the spatial resolution cannot be resolved and the results reported in this studied should be confirmed using more advanced eye tracking equipment. Second, as the text comprehension questions were limited in both number and diversity, results are limited with respect to the product of mental model building. Further research should aim to disentangle effects of different product-related mental model measures, such as summary writing, recall tasks, or differences between implicit and explicit questions. As Lorch and Lorch (1996) pointed out, headings might affect free recall tasks, and not summary writing. Thirdly, the results of this study do not answer the question as to why skipping probabilities moderate the effect of student-related characteristics on text comprehension. Future research should therefore focus especially on the eye movements of poor readers and focus on differences in mind wandering of these students (Nguyen, Binder, Nemier, & Ardoin, 2014).

In summary, this study shows in what way student- and text-related characteristics are associated with eye movements and how these factors influence text comprehension scores. The most important implication of this result is that less skilled readers should not only solely train reading speed when they want to become better comprehenders. Increasing reading speed for this group could also lead to poor comprehension scores and therefore these students should learn how to compensate for their lack of skill. Concluding, this study adds to the understanding of 4th graders' text comprehension by investigating both the process and product of reading comprehension in relation to student characteristics.

References

Andersson, R., Nyström, M., & Holmqvist, K. (2010). Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. *Journal of Eye Movement Research*, *3*(3), 6, 1-12.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Blanc, N., Kendeou, P., Van den Broek, P. W., & Brouillet, D. (2008). Updating situation models during reading of news reports: Evidence from empirical data and simulations. *Discourse Processes, 45*(2), 103-121. doi: 10.1080/01638530701792784.

Blythe, H. I., & Joseph, H. S.S.L. (2011).Children's eye-movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye-movements*. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199539789.013.0036.

Cain, K., Oakhill, J., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859. doi: 10.3758/BF03196414.

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

Calvo, M. G., Estevez, A., & Dowens, M. G. (2003). Time course of elaborative inferences in reading as a function of prior vocabulary knowledge. *Learning and Instruction*, *13*, 611–631. doi: 10.1016/S0959-4752(02)00055-5.

CED-Groep (2011). *NieuwsbegripXL.* Nederlands, Rotterdam: CED-Groep.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422-433. doi:10.3758/s13423-013-0565-2

De Leeuw, L., Segers, E., & Verhoeven, L. (2015). Role of text and student characteristics in real-time reading processes across the primary grades. *Journal of Research in Reading*.doi: 10.1111/1467-9817.12054.

Egberink, I. J. L., Janssen, N. A. M., & Vermeulen, C. S. M. (2009-2014). *COTAN Documentatie (www.cotandocumentatie.nl)*. Amsterdam: Boom test uitgevers.

Feenstra, H. (2008). *Begrijpend Lezen groep 6* [Reading Comprehension Grade 4]. Arnhem: Cito.

Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371-395. doi:10.1037/0033-295X.101.3.371.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language, 54*(3), 425-443. doi:10.1016/j.jml.2005.12.001.

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. doi: http://dx.doi.org/10.1016/j.visres.2009.09.012.

Hyönä, J., Lorch, R. F. Jr., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94,* 44-55. doi: 10.1037/0022-0663.94.1.44.

Jongen, I., & Krom, R. (2009). *DMT en AVI: AVI Handleiding* [DMT and AVI: AVI manual]. Arnhem: Cito.

Joseph, H., Nation, K., & Liversedge, S. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review, 42*(2), 207–222.

Juhasz, B. J., & Pollatsek, A. (2011). Lexical influences on eye movements in reading. In S. Liversedge, I. Gilchrist, & S. Eberling (Eds.) *Handbook on Eye movements, Oxford University Press*. doi: 10.1093/oxfordhb/9780199539789.013.0048.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4). 329-354. doi: http://dx.doi.org/10.1037/0033-295X.87.4.329.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122-149. doi:10.1037/0033-295X.99.1.122.

Kaakinen, J., Hyönä, 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*, 447–457. doi:10.1037/0278-7393.29.3.447.

Kaakinen , J. K., & Hyönä, J. (2007). Perspective effects in repeated reading: An eye movement study. *Memory & Cognition, 35*(6), 1323-1336. doi: 10.3758/BF03193604.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294-303. doi: 10.1037/0003-066X.49.4.294.

Kintsch, W. (2004). The Construction-Integration model of text comprehension and its implictions for instruction. In R. B. Ruddell, & N. J. Unrau, (Eds.), *Theoretical Models and Processes of Reading*. *Fifth Edition* (pp.1270-1328). Newark, USA: International Reading Association.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394. doi: 10.1037/0033-295X.85.5.363.

Kort, W., Schittekatte, M., Compaan, E. L., Bosman, M., Bleichrodt, N., Vermeir, G., Resing, W. C. M., & Verhaeghe, P. (2005). *WISC-III NL. Handleiding* [Manual]. *Nederlandse bewerking* [Dutch version]. London: The Psychological Corporation.

Lorch, R., & Lorch, E. (1996). Effects of Headings on Text Recall and Summarization. *Contemporary Educational Psychology, 21,* 261–278. doi: 0361-476X/96.

McMaster, K. L., Espin, C. A., & Van den Broek, P. (2014). The division for learning disabilities of the council for exceptional children making connections: Linking cognitive psychology and intervention research to improve comprehension of struggling readers. *Learning Disabilities Research & Practice*, *29*(1), 17–24. doi: 10.1111/ldrp.12026.

McNamara, D. S., & O'Reilly, T. (2009). Theories of comprehension and comprehension skill: Knowledge and strategies versus capacity and suppression. In A.M. Columbus (Ed.), *Advances in Psychology Research Vol. 62.* Hauppauge, NY: Nova Science Publishers.

Nguyen, K., Binder, K., Nemier, C., & Ardoin, S. (2014). Gotcha! Catching kids during mindless reading. *Scientific Studies of Reading*, *18*(4), 274-290. doi:10.1080/10888438.2013.876424.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior research methods, instruments, & computers*, *32*(3), 396-402. doi: 10.3758/BF03200807

Raven, J.C. (1960). *Guide to the standard progressive matrices: Sets A, B, C, D and E.* UK, London: H.K. Lewis & Co. Ltd.

Raven, J. C. (2006). *Raven: Standard progressive matrices (SPM) Handleiding.* Hardcourt Test Publishers.

Rayner, K. (1985). The role of eye-movements in learning to read and reading disability. *Remedial and Special Education, 6*(6), 53-60. doi: 10.1177/074193258500600609.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Pychological Bulletin, 124*(3), 372-422. doi: 10.1037/0033-2909.124.3.372.

Rayner, K., Kambe, G., & Duffy, S. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology, 53A*(4), 1061-1079. doi: 10.1080/713755934.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*, 445-526. doi: 10.1017/S0140525X03000104.

Reichle, E., Reineberg, A., & Schooler, J. (2010). Eye movements during mindless reading. *Psychological Science, 21*(9), 1300 –1310. doi: 10.1177/0956797610378686.

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition, 125*, 179–194. doi: http://dx.doi.org/10.1016/j.cognition.2012.07.004.

Schotter, E., 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. doi: 10.1177/0956797614531148.

Singer, M., Andrusiak, P., Reisdorf , P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition*, *20*(5), 539-548.

Swanson, H., Zheng, X., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: a selective meta-analysis of the literature. *Journal of Learning and Disabilities, 42*(3), 260-287.doi: 10.1177/0022219409331958.

Tellings, A., Hulsbosch, M., Vermeer, A., & Van den Bosch, A. (2014). BasiLex: an 11.5 million words corpus of Dutch texts written for children. *Computational Linguistics in the Netherlands Journal, 1*(4), 191-208.

Tiu, R. D. Jr., Thompson, L. A., & Lewis, B. A. (2003). The role of IQ in a component model of reading. *Journal of Learning Disabilities, 36*(5), 424-436. doi: 10.1177/00222194030360050401

Van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist approaches in accounts of reading comprehension. *Discourse Processes, 39,* 299-316. doi:10.1080/ 0163853X.2005.9651685.

Van den Broek, P., Virtue, S., Everson, M., Tzeng, Y., & Sung, Y. (2002). Comprehension and memory of science texts: Inferential processes and the construction of a mental representation. In J. Otero, J. Leon, & A.C. Graesser (Eds.), *The psychology of science textcomprehension* (pp. 131-154). Mahwah, NJ: Erlbaum.

Van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading. In H. Van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71-98). Mahwah, NJ: Erlbaum.

Van den Broek, P., Lorch, R.F. Jr, Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*, 1081-1087. doi: 10.3758/BF03206376.

Van den Broek, P. W., Risden K., & Husebye-Hartman E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In: R. F. Lorch Jr. & E. J. O'Brien (Eds.) *Sources of coherence in text comprehension*. Hillsdale, NJ: Erlbaum. 353-373.

Van der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. M. D. (2008). The Role of Two Reading Strategies in Text Comprehension: An eye fixation study in primary school children. *Journal in Research in Reading*, *31*(2), 203-223.

Verhoeven, L., & Perfetti, C. (2008). Advances in text comprehension: Model, process and development.*Applied Cognitive Psychology, 22*, 293–301. doi:10.1002/acp.1417.

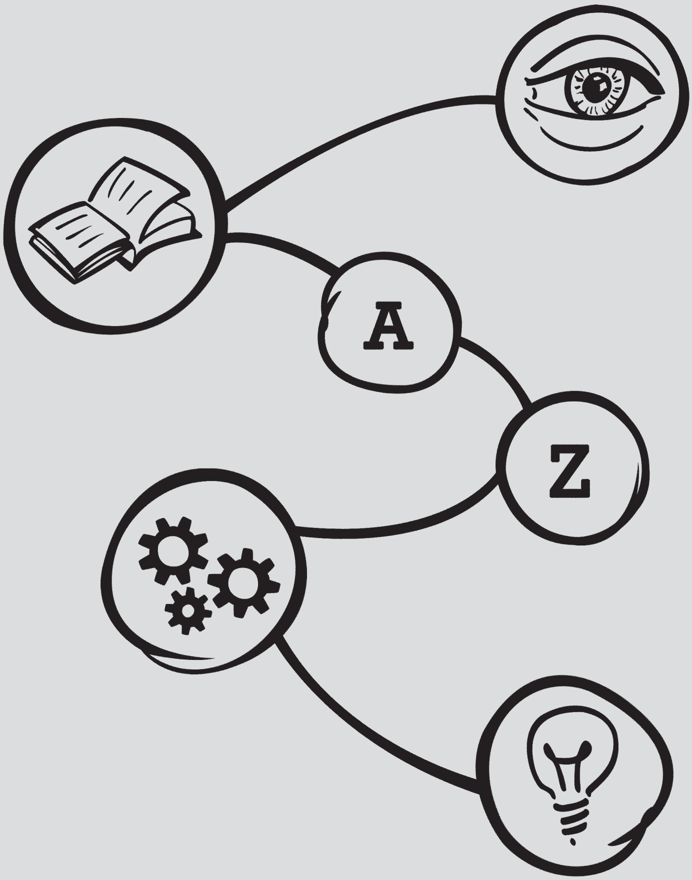
Verhoeven, L., & Vermeer, A. (1996). *Taalvaardigheid in de bovenbouw: Nederlands van autochtone en allochtone leerlingen in het basis en mlk-onderwijs.* Tilburg University Press.

Verhoeven, L., & Vermeer, A. (1999). *Leeswoordenschat* [Vocabulary for reading]. Arnhem: Cito.

Verhoeven, L., Keuning, J., Horsels, L., & Van Boxtel, L. (2013). *Testinstrumentarium Taal ontwikkelingsstoornissen voor Kinderen van 4 tot 10 Jaar.* Arnhem: Cito.

CHAPTER 4

Student- and text-related effects on real-time reading processes and reading comprehension in sixth graders[3]



[3]This chapter is based on:
De Leeuw, L., Segers, E., Fesel, S., & Verhoeven, L. (submitted). Student- and Text-related Effects on Real-Time Reading Processes and Reading Comprehension in Sixth Graders.

Abstract

The present study examined student- and text-related effects on real-time reading processes and reading comprehension in sixth graders. Sixty-three children read two expository texts consisting of ten paragraphs (each containing a heading and the remainder of the paragraph) nested within three sub-sections. First we examined effects of individual variation (decoding efficiency, vocabulary knowledge, short-term memory, working memory, reading comprehension skill, and non-verbal intelligence) on eye movements (processing times for heading and remainder). Second, we examined these effects in relation to reading outcomes (knowledge representation and multiple-choice questions). The results for heading showed longer processing times for deeper nested paragraphs. For the remainder of paragraph, processing times decreased throughout the text, leading to lower accuracy scores for questions concerning the end of the text for slower readers. Furthermore, individual differences in vocabulary were related to reading comprehension outcomes. It is concluded that both reading processes and individual variation contribute to reading comprehension in sixth graders.

Introduction

Readers need to build a coherent text representation during reading (Kintsch, 1994) in order to acquire knowledge from expository text. This results in a mental model that is stored in long-term memory (Van den Broek, 2012). Readers start creating a mental model from the onset of reading. During reading, this model is constantly updated by combining text elements stored in working memory. Proficient readers adapt their reading behaviour as the text unfolds (Bell, 2011). However, it is largely unknown to what extent developing readers adapt their reading behaviour while processing the text, and how this affects reading comprehension outcomes. Moreover, reading processes vary among readers (Blythe & Joseph, 2011; Rayner, 1985; Reichle, Rayner, & Pollatsek, 2003) due to linguistic and cognitive variation (e.g., variation in vocabulary knowledge). These individual differences are often not included in studies that focused on real-time reading processes of developing readers. Therefore, the present study investigated to what extent 6^{th} grade students' individual differences influence the time course of reading for comprehension and the quality of mental models when reading expository text.

Reading for comprehension

During text reading, readers construct a mental model by constantly updating the current model. They do this by adding new information to the information that is already gathered (Kintsch & Van Dijk, 1978; Van den Broek, Young, Tzeng, & Linderholm, 1999). Studies using think-aloud protocols showed mental model updating to occur after reading of each sentence, as students reported integrating arguments (either consistent or inconsistent with a previous statement) at clause boundaries (Blanc, Kendeou, Van den Broek, & Brouillet, 2008). Real-time reading studies investigating eye movements support these results and showed reading times to increase at sentence boundaries when reading science texts (Just & Carpenter, 1980).

Mental model updating in context has also been studied by investigating whether students detect inconsistencies within texts (Hyönä, Lorch, & Rinck, 2003, Van der Schoot, Reijntjes, & Van Lieshout, 2012). For example, texts are created in which in the beginning of a paragraph a vegetarian is introduced whom later would order a hamburger. These studies showed that readers detect such inconsistencies more easily when the vegetarian and the hamburger are presented closely together, compared to when it is presented further apart.

The research on mental model updating processes so far mainly focused on adults (Schotter, Tran, & Rayner, 2014). Studies on mental model updating in children reading texts are scare. This could be caused by the fact that methods traditionally used for studying reading processes, such as think-aloud protocols, require metacognitive skills, which are not fully matured in developing readers (Kuhn, 2000). The self-paced reading paradigm (Aaronson & Scarborough, 1976) is also a suboptimal method, because of its ecological validity. In such experiments, text is presented in segments and readers need to press a button to receive the next one. Pressing button highly interferes with the reading process, and hence the setup is not successful to mimic natural reading processes. With the introduction of more child friendly eye trackers, more studies have been conducted in children. These studies mainly focused on variation in processing different texts (Kaakinen, Hyönä, & Keenan, 2003; Van Silfhout, 2014), comparing skilled and less readers (Van der Schoot et al., 2012), or on task effects on comprehension (Kaakinen, Lehtola, & Paattilammi, 2015), but no studies focused on reading longer texts or individual variation among developing readers.

The process of mental model updating ultimately results in a mental model that is created during reading. This mental model consists of a 'network of propositions' (Kintch, 1994: 295) that improves with an increasing number of propositions and interconnections between propositions. After reading, this mental model is stored in long-term memory. At that point, a reader has learned from a text. However, not all information that is included in the mental model *during* reading is necessarily remembered *after* reading (Just & Carpenter, 1980; Kintsch, 1994). This is due to fact that some propositions are linked more directly to the main theme than others. These more directly linked propositions are recalled better after reading (Van den Broek et al., 1999; Van den Broek, Helder, & Van Leijenhorst, 2013).

A traditional way to measure reading comprehension is via the assessment of text comprehension questions. Explicit text comprehension questions measure the more basic, surface level of mental representation, while implicit questions tap into the deeper understanding of the text; the situation model (Kalamski, 2007; Kinstch, 2004). More recent methods to measure reading comprehension also tap into semantic relations of propositions within the text (Clariana, 2010). The semantic relations are examined using a related-judgment task, in which students are asked to rate the relatedness of word pairs that were selected from the text. Their judgments are compared to judgments of other readers in order to determine the quality of their mental model (see e.g., Fesel, Segers, Clariana, & Verhoeven, 2015).

Text- and student characteristics influence reading comprehension

Both text- and student characteristics influence the reading process. With regards to text characteristics, especially longer texts - that consist of several paragraphs - may pose challenges to mental model building. First, a reading needs to determine the hierarchy of the text. Information that is directly linked to the main topic is more important than other information. Also information may be more or less hierarchically salient. In longer texts, headings are particularly important, because they help the reader to structure the information within the text (McNamara, Ozuru, Best, & O'Reilly, 2007). The paragraph itself elaborates on the topic and provides examples. Headings thus help in creating new main nodes in the mental model with all information within the paragraph being connected to this node until a new section of text is introduced. For adults, it has been found that longer reading times of the headings are related to better subsequent performance on summary writing (Kaakinen et al., 2003).

Second, reading longer text requires the reader to adapt their reading process throughout the text in order to obtain optimal comprehension. Reading processes change from the beginning towards the end of the text; readers tend to increase their pace as the text proceeds (Linderholm, Virtue, Tzeng, & Van den Broek, 2004). There are two explanations for this increase in pace. First, crucial processes in mental model building occur at the beginning of the discourse, as readers have to get acquainted with the topic. In the beginning, they have to establish the main topic of the text before they can advance to thorough analysis and understanding of the text (Bell, 2011). As the text progresses, information updating becomes less effortful, because concepts may already be activated and prime upcoming information (Linderholm et al., 2004). An alternative approach to explain reading times to diminish throughout the text are reader fatigue (Graesser, Singer, & Trabasso, 1994; Van den Broek, Risden, & Husebye-Hartman, 1995) and mind wandering (Nguyen, Binder, Nemier, & Ardoin, 2014). Both readers fatigue and mind wandering cause readers to be less actively involved in reading, leading to a decrease in reading comprehension scores (Schad, Nuthmann, & Engbert, 2012; Nguyen et al., 2014).

Next to text characteristics, reading processes and reading outcomes also depend on student characteristics. With respect to reading processes, eye tracking studies evidenced that skilled readers have fewer (Rayner, 1985; Lester, Nagle, Johnson, & Fisher, 1979; McConkie, Zola, Grimes, Kerr, Bryant, & Wolff, 1991) and shorter fixations (McMaster, Espin, & Van den Broek, 2014) than poor readers. Furthermore, readers differ with respect to strategy behavior: adults writing good summaries paid relatively more attention to headings (Kaakinen et al., 2003) than adults that were poor at writing summaries. In a similar vein, skilled young readers paid more attention to important text elements than less skilled readers (Van der Schoot, Vasbinder, Horsley, & Van Lieshout, 2008). These results suggest that there is a positive relation between reading strategy behavior and text recall.

With respect to reading outcomes, text comprehension was found to be influenced by differences in linguistic and cognitive abilities among students. When readers have better vocabulary knowledge, this helps them to better understand the concepts within the text. And when readers are fluent decoders, they have more cognitive resources left for understanding. In other words, a better lexical quality of the words stored in the lexicon of the reader positively influences reading comprehension (Perfetti & Stafura, 2014). And in turn, better conceptual understanding increases the chance of memory-based inferences (Calvo, Estevez, & Dowens, 2003; Singer, Andrusiak, Reisdorf, & Black, 1992) on the one hand, and of making elaborate inferences (i.e., linking the text to prior knowledge) on the other hand (Van den Broek, Lorch, Linderholm, & Gustafson, 2001). Both lead to a better mental model. In other words, reading skill enhances memory for text (McMaster et al., 2014).

Finally, both short-term memory and working memory were found to be related to the amount of inferences that are generated during reading (Van den Broek et al., 2001), as well as to other measures of reading comprehension in both adults (Daneman & Merikle, 1996) and children (Cain, Oakhill, Barnes, & Bryant, 2001; Cain, Oakhill, & Bryant, 2004). Developing readers have to devote more cognitive resources to lower-level text processing (i.e. decoding, vocabulary) which limit the capacity available for higher-level text processing (Just & Carpenter, 1992).

The present study

The above presented overview from the literature shows that text- and student characteristics influence reading comprehension, and that with eye tracking studies, more insight is gained in online reading comprehension processes. However, to our knowledge, the impact of real-time text processing on comprehension outcomes in children has not been studied. Moreover, no studies yet focused on changes in eye movement behaviour as text progresses. This is especially relevant for developing readers, as variability across students is considerably large and could affect reading processes and outcomes.

Previous research has shown that both reading processes and reading outcomes are related to text and reader characteristics. However, most studies have been conducted on adult readers and reading processes of developing readers have received little attention (Blythe & Joseph, 2011). Moreover, the focus of prior research was either on reading processes or reading outcomes and few attempts have been to combine reading processes and outcomes in one design (Kaakinen et al., 2015; Schotter et al., 2014). Most importantly, none of these studies included intra- and inter individual variation among readers or texts. Therefore, the aim of this study was to examine the reading processes and reading outcomes in relation to text- and student characteristics in 6th grade students in the Netherlands. Two research questions were addressed:

1) Which processes underlie the reading of multi-paragraph expository text as a function of text structure and student characteristics?
2) To what extent do children's reading comprehension outcomes relate to their reading processes and student characteristics?

In order to answer these research questions students read two expository texts consisting of ten paragraphs each, starting with a heading and followed by the remainder of the paragraph. The structure of the text was simple: the first page introduced the topic (level 1). Thereafter three sections concerning different subtopics (level 2) each with two subthemes (level 3) were presented. After reading, students performed a related-judgment task (Clariana, 2010; Fesel et al., 2015) measuring the full mental model. In addition, students answered explicit and implicit multiple-choice questions. Mean reading times of the heading and remainder of the paragraph were related to text- (paragraph and section) and student-related characteristics (decoding, vocabulary, short-term memory, working memory, reading comprehension skill, non-verbal intelligence), as well as to text comprehension measures (knowledge representation and questions).

4

Method

Participants

Students from four 6th grade classes from three Dutch primary schools participated. From the 73 participants, 1 student was not included in the analyses because this student's reading comprehension score deviated more than 2,5 SD from the mean. Furthermore, 9 students were excluded due to missing eye movement data. The remaining 63 students (39 girls, 24 boys, M_{age}= 12;2 years, age range 11;3-13;5) were included in the analyses. Participants had a normal IQ, all scoring above the 25th percentile (M = 45.68, SD =5.48; *Standard Progressive Matrices;* Raven, 1960).

Materials

Target materials

Apparatus. To conduct this study, we used a Tobii T120 eye tracker with a sampling rate of 120 Hz. The distance between monitor and the head was approximately 70 cm. Participants were sitting on a chair adjusted to their height and the eye tracker was placed on a height adjustable table. Students were instructed to pull up their chair as tight as possible, put their chin on the chinrest and grab the mouse on the table. Texts were presented on a 17 inch screen with a 1280 x 1024 resolution with a white background and black letters. Text margins were 200 px from each side of the screen with font Arial, 30 px, line height 3 in normal style. The different heading levels were presented in a different font; level 1 headings (30 px, boldface), level 2 headings (25 px, normal) and level 3 heading (20 px, italicized).

Text material. Texts were all on geography topics: Oceania, Russia, South America and South Africa (Klois, Segers, & Verhoeven, 2013). All texts had a similar hierarchical structure as presented in Figure 1a: First the topic was introduced, followed by three main chapters each with two subchapters. The texts had a length of 10 pages with a mean length of 97.2 words per page (for Oceania: M = 101.4; Russia: M = 94.8; South America: M = 98.4; South Africa: M = 94.1) each starting with a heading (level 1, 2, or 3) with a mean length of 10.7 characters (for Oceania: M = 9.4; Russia: M = 10.5; South America: M = 11.4; South Africa: M = 11.5). An example paragraph is presented in Figure 1b. The participants clicked on either the left or right arrow on the screen to navigate back and forth. Text materials were identical for all students and the reading order of the topics was pseudo-randomly counterbalanced across participants.

a b

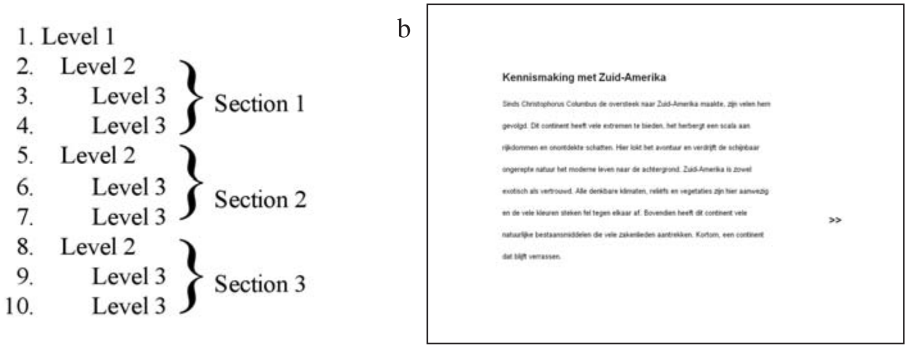


Figure 1. Overview of the target texts of this study (1a) and example page of the target text (1b).

Student-related characteristics

Decoding. Decoding skill was measured using a standardized word reading task for children between 7 and 12 years (*Een Minuut Toets* [One Minute Test], Van Brus & Voeten, 1973). On the card 116 words are presented, divided over four columns starting with one-syllable CVC words. Difficulty gradually increased to five syllables. Students were instructed to read aloud as many words as possible within one minute. Every correctly read word was scored as one point.

Vocabulary knowledge. Vocabulary knowledge was tested by a standardized passive vocabulary knowledge test (*Leeswoordenschattaak* [Vocabularytask], Verhoeven & Vermeer, 1999). This test consists of fifty multiple-choice items in which each word was presented in a short and uninformative context, e.g., ‘He sells vegetables’. The students were asked for the meaning of the underlined word. Four multiple choice options were presented including a synonym of the target word, e.g., ‘grass’, ‘green soup’, ‘salad’, and ‘edible plants’. Two practice words were discussed prior to testing. Questions regarding the task were answered, though no hints to answers were provided. Reported scores are the total number of correct answers with a maximum of 50.

4

Short-term memory (STM). Short-term memory was measured using a forward digit span memory task (*WISC-III^{NL}*, Kort et al., 2005). The researcher read aloud a string of digits using a falling intonation and pausing one second between the digits. The students were instructed to remember the digits in same order as presented. The strings started short (n = 2) with two attempts for each string length. Whenever children correctly remembered at least one of two strings, the researcher continued with a longer string, adding one digit until a maximum (n = 9) was reached. Each correctly remembered string accounted for one point with a maximum of 16.

Working memory (WM). Working memory was measured by a backward digit span memory task (*WISC-III^{NL}*, Kort et al., 2005). The researcher read aloud a string of digits using a falling intonation and pausing one second between the digits. The students were instructed to remember the digits in reversed order. The strings started short (n = 2) with two attempts for each string length. Whenever students correctly remembered at least one of the two strings, the researcher continued with a longer string, adding one digit until a maximum (n = 8) was reached. Each correctly remembered string accounted for one point with a maximum of 14.

Reading comprehension skills. Reading comprehension skills were measured using a standardized test for Grade 6 (Feenstra, 2008). This test consisted of two parts. The first part contained 25 questions and the second part consisted of 30 multiple choice questions. Questions were both aimed at sentence and text level, testing both reading comprehension skill at the local and global level. The second part was adapted to their reading level measured in the first part; poor readers received an easier version than the good readers. The scores were transformed into respective age norms (months of formal reading instruction) using item response theory models, which enables across test and across grade comparisons.

Non-verbal intelligence. To test the non-verbal intelligence of the students, a *Standard Progressive Matrices* (Raven, 1960) test was administered. This multiple-choice test consists of 60 items that increase in difficulty. For each item, the student needs to identify the missing element that completes the pattern shown in the question. Items are divided over five set (A, B, C, D, and E) with 12 items each. In set A and B, six answer options are presented, and in the other sets eight answers are provided. Prior to testing, the first and second item were discussed as an example. Every item was scores as a point and hence the maximum score was 60.

Dependent variables

Knowledge representations. To measure knowledge representations for each text, a related-judgment task was used (KU-mapper software, Clariana & Wallace, 2009; Taricani & Clariana, 2006). For each text, the 15 most important concept terms were selected by several proficient adult readers (Klois et al, 2013) based on frequency and meaningfulness in the content of an overall text comprehension. All possible pairs, in total 105, of these 15 words were randomly presented on a computer screen. Students were asked to judge the relatedness of these pairs by clicking on a scale ranging from 1 (*unrelated*) to 5 (*highly related*) (see Clariana & Wallace, 2009; Taricani & Clariana, 2006). Note that the original scale used for adults ranged from 1 to 9. We narrowed the scale in order to make the task more suitable for children. Instructions on the task were given in both written and oral form.

Next, a pathfinder scaling algorithm transforms the matrixes of children's ratings into network structures (see Goldsmith, Johnson, & Acton, 1991; Schvaneveldt, 1990; Trumpower, Sharara, & Goldsmith, 1991). The judgment of each participant and the resulting network structure can be compared to a referent network structure/model (Acton, Johnson, & Goldsmith, 1994; Gonyalvo, Cañas, & Bajo, 1994). To compare the children's knowledge representation to a non-sequential model, we calculated for each of the four text topics an average knowledge structure/model from pair-wise judgments of all participants. The similarity between two networks is determined by the correspondence of links in the two networks. The similarity is the intersection divided by the union number of links in the two networks. Two identical networks will yield a similarity of 1 and two networks that share no links will yield similarity of 0.

Text comprehension. To assess students' text comprehension, participants answered 20 multiple-choice questions on the text with four possible answers after they read the text (see Klois et al., 2013). Half of the questions were explicit questions (based on information that was explicitly stated in the text) and the other half were implicit questions (that needed to be inferred from the text. In order to improve reliability, several items were deleted for Oceania (n = 3), Russia (n = 1), South America (n = 5) and South Africa (n = 5). Internal consistency for each text was moderate to good, Oceania (α = .723), Russia (α = .745), South America (α = .645), and South Africa (α = .630).

Procedure

This study was part of a larger study testing the effect of strategy training. The procedure described here describes the two phases of the experiment that involved the study's pre test. In the first phase of the study, students were individually tested on decoding speed, short-term memory, and working memory. This session lasted about 15 minutes. Vocabulary knowledge, reading comprehension, and non-verbal intelligence tests were administered in class divided over several sessions. The vocabulary test lasted about 15 minutes. The reading comprehension test was administered in two sessions; one that lasted about 40 minutes and one of 50 minutes. Finally, the non-verbal intelligence test lasted about 45 minutes.

In the second part, we measured pre-test effects of eye movements individually in a quiet room. Students were positioned in front of the eye tracker, with their dominant hand on the mouse and their chin in a rest. Participants were instructed to read the texts for comprehension and that they would receive questions and assignments afterwards. All instructions were read aloud by the instructor and students were asked to read along. The instructions consisted of three pages in which students learned how to navigate back and forth. After instruction, the calibration procedure was started using nine red fixation dots on a white background. Instruction and reading lasted about 15 minutes. Subsequent to reading, they performed several tasks in another quiet room. The first task was to the related-judgment task, which lasted about 10 minutes. Secondly, students answered 20 multiple choice questions, which lasted another 10 minutes.

Data analyses

In order to analyze the eye movement data, all headings and the remainder of the paragraph were considered to be different areas of interest (AOI). Thereafter, total reading time of the heading and remainder of the paragraph were calculated per word and log transformed. Reading times that were 2.5 SDs from the mean were deleted from further analyses.

To determine the role of student and text-related characteristics on reading times of the heading and the remainder of the paragraph, we conducted linear mixed effects regression models (*LMER*; Baayen, 2008) with the total reading time measures (heading and remainder of the paragraph) as dependent variables. A backward stepwise selection procedure was used, deleting all predictors and interactions that did not reach significance at the level of 5% (LMER; Baayen, 2008). The full model contained main effects of text-related characteristics paragraph and section, as well as interactions among these variables. Also, main effects of the student-related characteristics decoding, vocabulary, short-term memory, working memory, reading comprehension skill and non-verbal intelligence. Random effects included in the model were participant, text, and question.

For the similarity data, a linear mixed effects regression model was run, including main fixed effects of page, level, and total reading times (heading and remainder of the paragraph), as well as interactions among these variables. Furthermore, student-related characteristics were included: decoding, vocabulary, short-term memory, working memory, reading comprehension skill and non-verbal intelligence. Random effects that were included were participant and text. For the text comprehension questions, a logit mixed effects regression model was constructed including answers on the comprehension questions. Similar effects as described for the similarity data were tested, including random effects of question and explicitness of the question.

Finally, forward model comparisons - of the reduced and unreduced models - based on log-likelihood ratio tests were conducted to determine the maximum random slope effect structure for each model. Thereafter, the fitted model was re-examined and insignificant fixed effects were deleted. For mixed linear-effect models and mixed logit models, respectively t-values and z-values are reported.

Results

Descriptives

In Table 1, means and SD's of decoding, vocabulary knowledge, short-term memory, working memory, reading comprehension skill and non-verbal intelligence are reported. Most of the student-related characteristics were found to correlate moderately to strong among each other, but vocabulary was not correlated to decoding and memory measures. Mean reading times on the heading and remainder of the paragraph for each paragraph are presented in Table 2 and Figure 2. Table 2 also indicates in what way the sections and headings are distributed throughout the pages.

Table 1

*Correlations Mean Scores and Standard Deviations of Student-related Characteristics (*N*=63)*

Student characteristics	1	2	3	4	5	*M*	*SD*	*Range*
1. Decoding	-					75.81	16.38	40-111
2. Vocabulary	0.196	-				42.40	4.34	30-48
3. Short-term memory	0.471**	0.146	-			8.57	1.62	6-13
4. Working memory	0.321*	0.197	0.414**	-		5.86	1.73	2-11
5. Reading comprehension	0.461**	0.426**	0.415**	0.349**	-	59.32	19.02	23-119
6. Non-verbal intelligence	0.162	0.443**	0.318*	0.334**	0.552**	45.68	5.48	33-56

Note. $^{**}p < 0.05$. $^{*}p < 0.001$

4

Table 2

*Means and Standard Deviations of Reading Time Durations per Word in ms per Paragraph (*N*=63).*

			Heading	Remainder
Paragraph	Section	Level	*M (SD)*	*M (SD)*
1	-	1	1084 (586)	393(168)
2	1	2	694 (371)	383 (148)
3	1	3	571 (321)	415 (160)
4	1	3	948 (434)	331 (135)
5	2	2	447 (255)	371 (150)
6	2	3	514 (332)	348 (152)
7	2	3	708 (342)	344 (163)
8	3	2	890 (434)	315 (149)
9	3	3	991 (486)	384 (172)
10	3	3	980 (415)	267 (145)

Note. Means and SDs are calculated on aggregated means per participant.

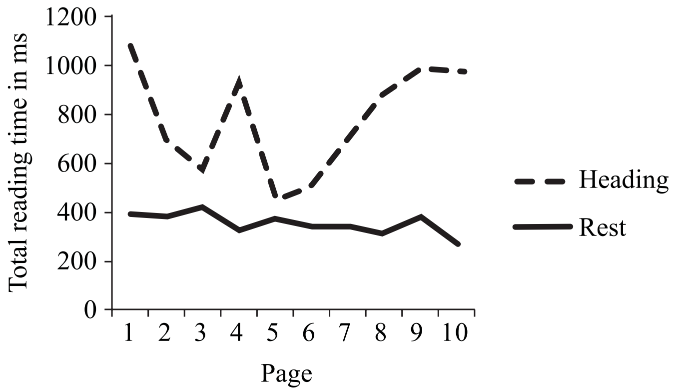


Figure 2. Reading times per word for the heading and the remainder of the paragraph.

Effects of text- and student-related characteristics on reading processes

In order to answer the first research question, a linear mixed regression effect model (Baayen, 2008) was run with the reading time as dependent variable and text-related characteristics (paragraph number and section) and individual differences (decoding, vocabulary, short-term memory, working memory, reading comprehension skill, non-verbal intelligence) as predictors. As random variables, participant, text, and question were included, as well as their random slopes. The total number of trials included in this analysis is 2032.

Reading times of the headings

The results showed negative main effects of decoding, $ß$ = -0.140, t = -3.248, p =.001, and section, $ß$ = -1.028, t = -5.115, p <.001, but not for paragraph, $ß$ = -0.051, t = -0.612, p = .541 (Table 3). With respect to decoding, this indicates that students with lower scores on decoding had longer reading times. For paragraph number, reading times on the heading were not longer for paragraphs that are presented at the beginning of the text, compared to the end of the text. However, within section, the reading times of the heading of the first section was longer (M =1067 ms) than the reading time for the heading of the other sections ($M_{section1}$ = 742 ms, $M_{section2}$ =555 ms, $M_{section3}$ = 944 ms). Finally, there was an interaction between paragraph number and section, $ß$ = 0.115, t = 5.665, p < .001. Further exploration of this interaction showed that reading times of headings across paragraphs increased within the first, $ß$ = 0.245, t = 2.044, p = .041, and second section, $ß$ = 0.231, t = 4.665, p < .001, but not in the third section, $ß$ = 0.067, t = 1.348, p = .178.

Table 3

Results on the Effects of Text and Student-Related Characteristics on Total Reading Times of the Heading

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	0.718	3.341	< .001
Student-related characteristics			
Decoding	-0.140	-3.248	= .001
Text-related characteristics			
Paragraph	-0.051	-0.612	= .541
Section	-1.028	-5.115	< .001
Section: paragraph	0.115	5.665	< .001
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.194	199.84	< .001
Participant: paragraph	0.035	25.52	< .001
Participant: section	0.334	26.82	< .001
Text	0.052	12.89	< .01
Text: section	0.010	12.37	< .01
Question	0.038	51.39	< .001

Reading times of the remainder of the paragraph

Table 4 present the results on analysis of text- and student-related characteristics on the remainder of the paragraph. As for the headings, the results showed a main effect of decoding, $ß$ = -0.585, t = -6.459, p < .001, and section, $ß$ = -0.067, t = -3.725, p < .001. The main effect of decoding indicates that lower decoding skills lead to slower reading times. The effect of section evidenced longer reading times of the paragraphs in the beginning of the text (M = 393) compared to those at the end of the text (M = 267). There were no significant interactions. For a graphical overview of the results, the mean reading times of the heading and remainder of the paragraph per word are displayed in Figure 2.

Table 4

Results on the Effects of Text and Student-Related Characteristics on Total Reading Times of the Remainder of the Paragraph

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	0.457	3.008	= .003
Student-related characteristics			
Decoding	-0.585	-6.459	< .001
Text-related characteristics			
Paragraph	-0.067	-3.725	< .001
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.589	1413.6	< .001
Participant: paragraph	0.004	81.71	< .001
Text	0.013	47.58	< .01
Text: decoding	0.003	9.88	= .007
Question	0.044	206.2	< .001
Question: decoding	0.001	8.45	= .015

Effects of text- , student- and process-related characteristics on text comprehension

Knowledge representation

To answer the second research question, on whether individual differences predict the mental representation after reading, we first analysed the student-related effects on students' knowledge representations. A linear mixed regression model was run with student-related characteristics (decoding, vocabulary knowledge, short-term memory, working memory, reading comprehension skill and non-verbal intelligence) as fixed variables, as well as reading times of the heading and the remainder of the paragraph. Furthermore, interactions with paragraph were included. Participant, text and question were included as random variables. The total number of trials was 1974 and results are reported in Table 5. As can be deducted from Table 5, the knowledge representation was solely predicted by vocabulary knowledge, $ß = 0.027$, $t = 3.329$, $p < .001$, indicating that children with higher vocabulary were scoring better at knowledge representations. There were no other significant main or interaction effects.

Table 5

Results on the Effects of Text and Student-Related Characteristics and Reading Times on Knowledge Presentations

Predictor: Fixed effects	*ß*	*t*	*p*
Intercept	0.207	20.831	< .001
Student-related characteristics			
Vocabulary	0.027	3.329	< .001
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.002	2193	< .001
Text	< 0.001	125.83	< .001
Text: vocabulary	< 0.001	48.56	< .001

Text comprehension questions

To examine effects of text- and student-related characteristics on answering text comprehension questions, a logit mixed regression model analysis on the questions of the text was conducted with (in)correct answers as a dichotomous dependent variable. A similar model was defined as for the reading times, and included student-related characteristics (decoding, vocabulary knowledge, short-term memory, working memory, reading comprehension skill) as fixed variables. In addition, readings times of the heading and the remainder of the paragraph were included, as well as interactions with paragraph number. Participant, text, and question were added as random variables. The total number of trials was 1995.

The results in Table 6 showed main effects of vocabulary, $ß = 0.361$, $z = 5.960$, $p < .001$, a marginal effect for explicitness of the questions, $ß = -0.241$, $z = -1.908$, $p = .056$, but no main effects for paragraph number and reading times of the remainder of the paragraph (*p*s > .157). The main effect of vocabulary showed that students with higher vocabulary knowledge were better at answering text comprehension questions. The explicitness-trend indicated that explicit questions were answered correctly more frequently than implicit questions.

Moreover, an interaction effect of the remainder of the paragraph and paragraph number was found for reading times, $ß = 0.033$, $z = 2.007$, $p = .048$. To further explore this interaction, analyses were rerun for the 50% faster and the 50% slower reading times. The effect of paragraph was significant for slow reading times, $ß = -0.054$, $z = -2.153$, $p = .003$, but not for fast reading times, $ß = 0.043$, $z = 1.558$, $p = .119$. When reading times are shorter, the position of the text does not influence answering text comprehension questions, whereas for longer reading times questions are more frequently answered correctly when they concern information that was presented at the beginning of the text. A graphical representation of this effect can be found in Figure 3.

Table 6

Results on the Effects of Text and Student-Related Characteristics and Reading Times on Text Comprehension Questions

Predictor: Fixed effects	*ß*	*z*	*p*
Intercept	0.186	0.738	= .461
Student-related characteristics			
Vocabulary	0.361	5.960	< .001
Reading times			
Remainder of the paragraph	-0.151	-1.415	= .157
Paragraph	-0.004	-0.171	= .865
Remainder of the paragraph: paragraph number	0.033	2.007	= .048
Question explicitness	-0.241	-1.908	= .056
Predictor: Random effects	*Explained variance*	χ^2	*p*
Participant	0.077	7.268	= .007
Text	0.037	6.305	= .012
Question	0.001	4.458	= .038

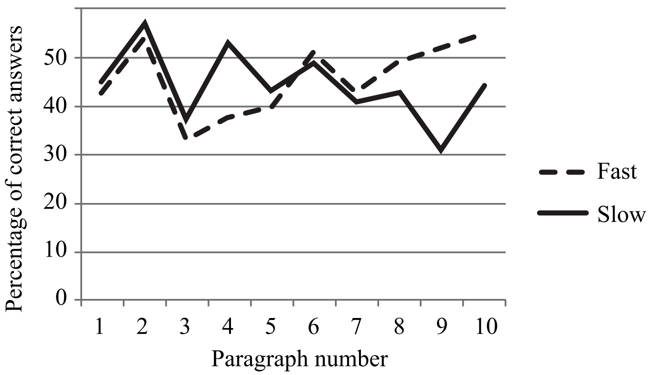


Figure 3. Scores on text comprehension questions for each paragraph for fast and slow reading.

4

Discussion

The present study examined the relation of the time course and text comprehension of reading long expository texts and effects of individual differences among 6th grade students. The results showed that the reading times increased for headings that were hierarchically more salient. Reading times of the remainder of the paragraph declined towards the end of the text. As expected, higher decoding skills were related to shorter reading times. Furthermore, shorter reading times of the remainder of the paragraph (not the heading) did not influence knowledge representation as tested by a related-judgment task (Clariana, 2010), but did have a negative effect on answering text comprehension questions. Finally, children who were slower readers answered text comprehension questions better when the questions concerned paragraphs at the beginning of the text.

With respect to the first research question, the results confirm that developing readers spend more time on orientation when starting a new text just as adults do (Bell, 2011). This supports the idea that building a mental model is more effortful at the beginning. A reader needs to explore the topic and get a first gesture of the text. Thereafter, reading becomes easier as the text unfolds (Bell, 2011; Linderholm et al., 2004). Hence connecting the propositions within a text cannot considered to be evenly effortful across longer text reading. An important remark is that our experimental design did not directly compare an identical target section at the beginning and end of a text. The beginning and end did not differ in sentence length or word frequency (often related to text difficulty), though this does not rule out effects of differences between text segments on comprehension processes and scores.

The finding that deeper-structured headings results in longer reading times indicates that it is more difficult to link these deeper structured headings to the mental model (Van den Broek et al, 2013). This would indicate that links that are further away from the main topic are harder to process. The effect of hierarchical structure becomes smaller as the text progresses, suggesting that students get acquainted to the text structure. This latter observation can also be linked to the faster overall processing effect we found. It seems as if students get acquainted with a topic or the structure of the text.

The second research question involved the relation of children's reading comprehension outcomes, their reading processes and student characteristics. Two comprehension outcomes were examined: related-judgment task, comprehension questions. With respect to knowledge representations as measured by the *related-judgment task* (Clariana, 2010), the results indicated that the reading process did not predict reading outcomes. However, student characteristics were found to be related; higher vocabulary knowledge was related to higher knowledge representations. The importance of vocabulary knowledge is in line with the lexical quality hypothesis (Perfetti & Stafura, 2014), but also in line with other studies that showed the importance of vocabulary knowledge in reading comprehension (Calvo et al., 2003; Singer et al., 1992).

With regard to reading *comprehension questions*, we found that accuracy scores depend on the position of the information in the text when reading slow reading, not for fast reading. Post–hoc analyses showed that text reading times were related to decoding efficiency. Whenever readers progress through a text, faster readers may benefit from a well-established mental model (Bell, 2011), leading to faster reading times. However, slow readers might not benefit from a good mental model, which seems to lead to a lower-quality mental model. The questions that pertain to parts at the end of the text are hence more difficult for these students. The less advanced mental model might be due to a lower-quality mental model due to a lack of skill, i.e., they are not good in disentangling (un)important information (McNamara & O'Reilly, 2009), leading to a too elaborate mental model that is difficult to keep updated. Also, slow readers may suffer from reader fatigue (Graesser, et al., 1994; Van den Broek et al., 1995) or mind wandering (Nguyen et al., 2014) leading to worst mental models.

At this point, several limitations should be addressed. First, the time course of reading longer expository texts was solely investigated with respect to global eye movement measures. Especially fine-grained effects within the remainder of the paragraph could be very informative, for example when looking at the first and final sentence of the paragraph. The set up of this experiment was not designed to perform such detailed analyses, since the length of sentences varied too much. Another limitation is related to word frequency effects. Since the texts in this study were designed to introduce new topics, allowing students to acquire new information, words within the text, especially the heading, were low in frequency hindering us to include frequency effects. Future studies should address the effects of word frequency by controlling for frequency or by comparing high and low word frequency in headings and paragraphs.

To conclude, this study showed differences in real-time reading processes with respect to reader-related and text-related characteristics. Reading comprehension processes rely on individual differences in decoding efficiency, while reading comprehension outcomes rely on vocabulary knowledge. Developing readers change their reading behaviour as they progress through the text and increase their pace towards the end to the text. For students with high decoding skills this increase in pace does not result in lower comprehension scores on questions concerning the information that was presented at the end. However, students with poor decoding skills showed lower comprehension scores on question concerning the final paragraphs of the text. In all, these results suggest that reading comprehension may be influenced via adapting text structure to the individual needs of the reader.

References

Aaronson, D., & Scarborough, H. S. (1976). Performance theories for sentence coding: some quantitative evidence. *Journal of Experimental Psychology: Human Perception and Performance, 2*(1), 56-70. doi: 10.1016/S0022-5371(77)80052-2.

Acton, W. H., Johnson, P. J., & Goldsmith, T. E. (1994). Structural knowledge assessment: Comparison of referent structures. *Journal of Educational Psychology, 86*(2), 303-311. doi:10.1037/0022-0663.86.2.303.

Baayen, R. H. (2008). *Analyzing Linguistic Data: a Practical Introduction to Statistics using R.* Cambridge: Cambridge University Press.

Bell, A. (2011). Re-constructing Babel: Discourse analysis, hermeneutics and the Interpretive Arc. *Discourse Studies, 13*(5), 519-568. doi:10.1177/ 1461445611412699.

Blanc, N., Kendeou, P., Van den Broek, P. W., & Brouillet, D. (2008). Updating situation models during reading of news reports: Evidence from empirical data and simulations. *Discourse Processes, 45*(2), 103-121. doi: 10.1080/ 01638530701792784.

Blythe, H., & Joseph, H. (2011). Children's eye-movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds), *The Oxford Handbook of Eye-movements* (pp.). doi:10.1093/ oxfordhb/9780199539789.013.0036.

Cain, K., Oakhill, J., Barnes, M.A., & Bryant, P. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859. doi:10.3758/ BF03196414.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31-42. doi:10.1037/0022-0663. 96.1.31.

Calvo, M. G., Estevez, A., & Dowens, M. G. (2003). Time course of elaborative inferences in reading as a function of prior vocabulary knowledge. *Learning and Instruction*, *13*, 611–631. doi:10.1016/S0959-4752(02)00055-5.

Clariana, R. B. (2010). Deriving individual and group knowledge structure from network diagrams and from essays. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-Based Diagnostics and Systematic Analysis of Knowledge*. Springer US.

Clariana, R. B., & Wallace, P. E. (2009). A comparison of pair-wise, list-wise, and clustering approaches for eliciting structural knowledge in information systems courses. *International Journal of Instructional Media*, *36*(3), 287-302. doi: 10.2190/2287-3204-06HG-2R37.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4). 422-433. doi:10.3758/ s13423-013-0565-2.

4

Feenstra, H. (2008). *Begrijpend Lezen groep 6* [Reading Comprehension Grade 4]. Arnhem: Cito.

Fesel, S., Segers, E., Clariana, R. B., & Verhoeven, L. (2015). Quality of children's knowledge representations in digital text comprehension: Evidence from pathfinder networks. *Computers in Human Behavior, 48,*135-146. doi:10.1016/ j.chb.2015.01.014.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 83*(1), 88-96. doi:10.1037/0022-0663.83.1.88.

Gonzalvo, P., Cañas, J. J., & Bajo, M. (1994). Structural representations in knowledge acquisition, *Journal of Educational Psychology, 86*(4), 601-616. doi:10.1037/ 0022-0663.86.4.601.

Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371-395. doi:10.1037/ 0033-295X.101.3.371.

Hyönä, J., Lorch, R. F., Jr., & Rinck, M. (2003). Eye-movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye-movement research* (pp. 313-334). Amsterdam: Elsevier Science.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, *87*(4), 329-354. doi:10.1037/0033-295X.87.4.329.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122-149. doi:10.1037/0033-295X.99.1.122.

Kaakinen, J., Hyönä, 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*, 447-457. doi:10.1037/0278-7393.29.3.447.

Kaakinen, J. K., Lehtola, A., & Paattilammi, S. (2015). The influence of a reading task on children's eye movements during reading. *Journal of Cognitive Psychology.* doi: 10.1080/ 20445911. 2015.1005623.

Kamalski, J. (2007). *Coherence marking, comprehension and persuasion: On the processing and representation of discourse*. Utrecht: LOT.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294-303. doi:10.1037/0003-066X.49.4.294.

Kintsch, W. (2004). The Construction-Integration model of text comprehension and its implictions for instruction. In: R. B. Ruddell, & N. J. Unrau, (Eds.), *Theoretical models and processes of reading*. *Fifth edition* (pp.1270-1328). Newark, USA: International Reading Association.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394. doi:10.1037/0033-295X.85.5.363.

Klois, S. S., Segers, E., & Verhoeven, L. (2013). How hypertext fosters children's knowledge acquisition: The roles of text structure and graphical overview. *Computers in Human Behavior, 29,* 2047-2057. doi:10,1016/j.chb.2013.03.013.

Kort, W., Schittekatte, M., Compaan, E. L., Bosman, M., Bleichrodt, N., Vermeir, G., Resing, W. C. M., & Verhaeghe, P. (2005). *WISC-III NL. Handleiding* [Manual]. *Nederlandse bewerking* [Dutch version]. London: The Psychological Corporation.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*(5), 178-181. doi:10.1111/1467-8721.00088.

Lester, A. L., Nagle, R. J., Johnson, G., & Fisher, D. (1979). Eye-movements dynamics of good and poor readers then and now. *Journal of Reading Behavior*, *9*(4), 319-328. doi:10.1080/10862967909547338.

Linderholm, T., Virtue, S., Tzeng, Y., & Van den Broek, P. (2004). Fluctuations in the availability of information during reading: Capturing cognitive processes using the Landscape Model. *Discourse Processes, 37,* 165-186. doi:10.1207/ s15326950dp3702_5

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye-movement during reading. In J. F. Stein (Ed.), *Vision and visual dyslexia* (pp. 251-262). London: Macmillan Press.

McMaster, K. L., Espin, C. A., & Van den Broek, P. (2014). Making connections: Linking cognitive psychology and intervention research to improve comprehension of struggling readers. *Learning Disabilities Research & Practice*, *29*(1), 17-24. doi: 10.1111/ldrp.12026.

McNamara, D. S., & O'Reilly, T. (2009). Theories of comprehension and comprehension skill: Knowledge and strategies versus capacity and suppression. In A. M. Columbus (Ed.), *Advances in Psychology Research, Vol. 62*. Hauppauge, NY: Nova Science Publishers.

McNamara, D., Ozuru, Y., Best, R., & O'Reilly, T. (2007). The 4-pronged comprehension strategy framework. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 465-496). Mahwah, NJ: Erlbaum.

Nguyen, K., Binder, K., Nemier, C., & Ardoin, S. (2014). Gotcha! Catching kids during mindless reading, *Scientific Studies of Reading*, *18*(4), 274-290. doi:10.1080/ 10888438.2013.876424.

Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*, 22-37. doi:10.1080/ 10888438.2013. 827687.

Raven, J. C. (1960). *Guide to the standard progressive matrices: Sets A, B, C, D and E.* UK, London: H.K. Lewis & Co. Ltd.

Rayner, K. (1985). The role of eye-movements in learning to read and reading disability. *Remedial and Special Education, 6*(6), 53-60. doi:10.1177/ 074193258500600609.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*, 445-526. doi:10.1017/S0140525X03000104.

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition,125*, 179-194. doi:10.1016/j.cognition.2012.07.004.

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*, 1218-1226.

Schvaneveldt, R. W. (Ed.) (1990). *Pathfinder associative networks: Studies in knowledge organization.* Norwood, NJ: Ablex.

Singer, M., Andrusiak, P., Reisdorf, P., & Block, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition*, *20*(5), 539-548.

Taricani, E., & Clariana, R. B. (2006). A technique for automatically scoring open-ended concept maps. *Educational Technology Research and Development*, *54*(1), 65-82. doi:10.1007/s11423-006-6497-z .

Trumpower, D. L., Sharara, H., & Goldsmith, T. E. (2010). Specificity of structural assessment of knowledge. *The Journal of Technology, Learning, and Assessment, 8*(5). Received from http://files.eric.ed.gov/fulltext/EJ873674.pdf

Van Brus, B., & Voeten, M. (1973). *Een-minuuttest [One minute test].* Nijmegen, Berkhout.

Van den Broek, P. (2012). Individual and developmental differences in reading comprehension: Assessing cognitive processes and outcomes. In J. P. Sabatini, E. R. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 39-58). Lanham: Rowman & Littlefield Education.

Van den Broek, P. W., Helder, A., & Van Leijenhorst, L. (2013). Sensitivity to structural centrality: Developmental and individual differences in reading comprehension skills. In M. A. Britt, S. R. Goldman, & J. F. Rouet (Eds.), *Reading: From Words to Multiple Texts* (pp. 132-146). New York: Routledge.

Van den Broek, P., Lorch, R. F. Jr, Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*, 1081-1087. doi:10.3758/BF03206376.

Van den Broek, P. W., Risden K., & Husebye-Hartman E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch Jr. & E. J. O'Brien (Eds.) *Sources of coherence in text comprehension* (pp. 353-373). Hillsdale, NJ: Erlbaum.

Van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading. In H. Van Oostendorp, & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71-98). Mahwah, NJ: Erlbaum.

Van de 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. doi:10.1007/s11145-011-9337-4.

Van der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. M. D. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal in Research in Reading*, *31*(2), 203-223. doi:10.1111/j.1467-9817.2007.00354.x.

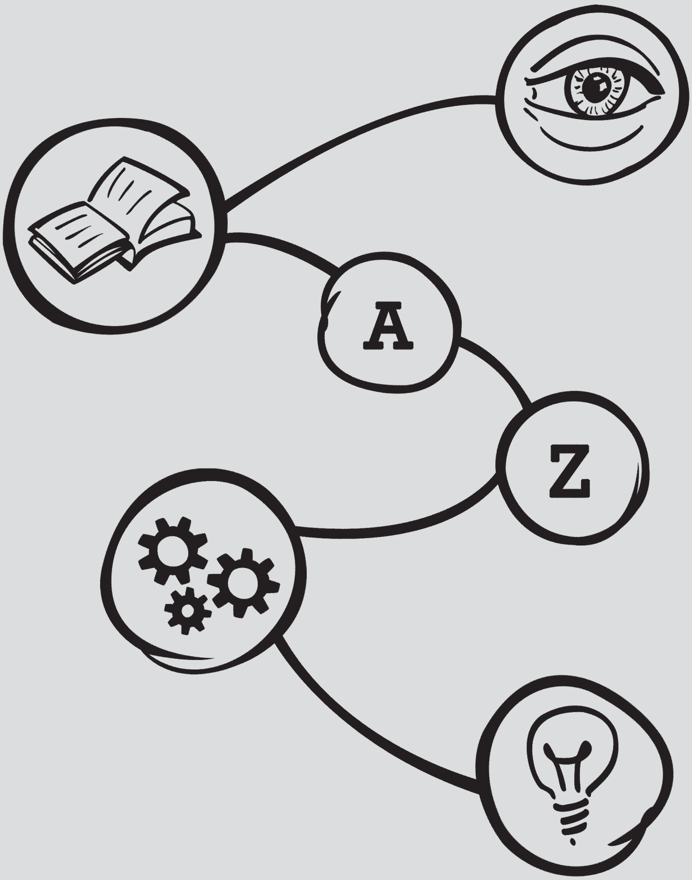
Van Silfhout, G. (2014). *Fun to read or easy to understand? Establishing effective text features for educational texts on the basis of processing and comprehension research.* Utrecht: LOT.

Verhoeven, L., & Vermeer, A. (1999). *Leeswoordenschat* [Vocabulary for reading]. Arnhem: Cito.

4

CHAPTER 5

Context, task, and reader effects in children's incidental word learning from text[4]



[4] This paper has been published as:
Leeuw, L. de, Segers, E., & Verhoeven, L. (2014). Context, task, and reader effects in children's incidental word learning from text. *International Journal of Disability, Development and Education, 61*(3), 275-287. doi: 10.1080/1034912X.2014.934010

Abstract

Incidental word learning is influenced by context, task and reader characteristics. The present study aimed to determine the contribution and interactions of these factors for fifth grade students. The focus was on contextual differences: words meanings are inferred from local or global contexts. This effect was tested as a function of task: gap filling, inference questions and summary writing in comparison to the single reading of the text. Regarding the reader, the contribution of general vocabulary knowledge and working memory was determined. The results showed that words are better learned in local than global contexts, and that the higher-level tasks (inference questions and summary writing) enhanced word learning beyond the single reading of the text, whereas gap filling did not. General vocabulary knowledge was related to overall incidental word learning from text, whereas WM contributed to vocabulary gain from answering inference questions. It can be concluded that incidental word learning from text is optimal in local contexts, when doing higher-level tasks and when general vocabulary is high.

Introduction

The development of vocabulary is key to the extension of knowledge foundation. When children are at a certain reading level, reading materials are an important source of new vocabulary (Nagy & Scott, 2000). By making inferences from context, children form hypotheses about the meaning of each newly encountered word (Cain, Lemmon, & Oakhill, 2004; Fukkink, 2005). As the amount of reading material increases for beginning readers, incidental word learning becomes a more important skill. But spontaneous inference generation - needed for incidental word learning - appears to be dependent of factors related to context, task, and reader. With respect to context, inferences from local contexts were found to be easier than inferences from larger, global contexts (Bolger, Balass, Landen, & Perfetti, 2008). With regard to task, it was found that tasks which involve deeper text processing may lead to more and better inference generation (Van den Broek, Lorch, Linderholm, & Gustafson, 2001). With reference to the reader, it was evidenced that characteristics such as vocabulary knowledge and working memory capacity influence inference generation (McNamara & O'Reilly, 2009). In previous research, however, no attempt has been made to study the impact of all three factors on incidental word learning in one design. Therefore, the present study aimed to shed more light on this issue by comparing word learning for which the context required the reader to make either local or global inferences in three different types of text processing tasks, in perspective of the children's vocabulary knowledge and WM abilities.

Context effects

The type of context of a new word influences if and how the new word meaning is learned. When a new word is encountered, it must be linked to the proper contextual clues. When the contextual cue is near - i.e. precedes or follows the word - the reader needs to make a *local inference.* We refer to contexts in which the contextual cue is near and local inferences are needed as local contexts. When the meaning needs to be extracted from a larger context this requires a *global inference* and this is what we refer to as global contexts. In other words, inferring the correct meaning of a newly encountered word relies on whether coherence is global or local. Consider the text segment in (1) and (2).

(1) Uit onderzoek blijkt dat sommige soorten vleermuizen antistoffen tegen ebola in zich hebben. Antistoffen zijn extra ridders die weten hoe ze een bepaald virus moeten verslaan.
‘Research has shown that some types of bats have antibodies against the ebola fever. Antibodies are extra knights who know how to defeat a certain virus.’

(2) Daarom zorgt de regering ervoor dat kippen niet in aanraking kunnen komen met trekvogels. Ze mogen niet naar buiten lopen. Maar ook met de ophokplicht sterven er kippen.
‘That is why the government makes sure chickens cannot contact migratory birds. They cannot go outside. But when poultry is kept indoors chicken die as well.’

To determine the meaning of the word *antistof* (antibody), a local inference needs to be made; the explanation that antibodies are extra knights is presented in the sentence after which the word was firstly introduced. Determining the meaning of the word *ophokplicht* (keep poultry indoors) is more difficult because the information is in a larger context. A global inference needs to be made; chicken will stay indoor because of this and therefore cannot contact migratory birds.

Global inferences are presumed to be more costly than local inferences. They require the reader to remember a larger segment of the text, whereas local inferences rely on the memory of smaller text segments. This cost-effect was shown for experienced readers in a study on bridging inferences by Singer, Andrusiak, Reisdorf, and Block (1992). Local inferences were found to be generated faster than global inferences. For word learning, cost-effects were found as well (Carnine, Kameenui, & Coyle, 1984; Swanborn & De Glopper, 1999). Carnine et al. (1984) found that when the contextual clues were close, word learning was enhanced for less experienced readers. Swanborn and De Glopper (1999) found, in a meta-analysis, that word learning from larger contexts is more difficult compared to smaller context.

Task effects

Reading comprehension tasks are often designed to improve text understanding. In order to understand the text, it is important for the reader to create a coherent text representation. Kintsch (2004) proposed three different text level representations that distinguish between poor and good representations. First, one needs to understand the sentences within the text (surface code), and second, the coherence between the independent sentences and segments (text base). Finally, the reader has to integrate the acquired knowledge with prior knowledge (situation model).

In order to construct higher levels of text representations, the reader does not only have to understand the separate sentences, but also needs to link the information of several text segments in order to make inferences. The quality of the text representation highly depends on the quantity and quality of the inferences that are made (Van den Broek et al., 2001, Tarchi, 2010): Local inferences construct shallow text representations, whereas global inferences, which are drawn across larger text segments, construct deeper text representations. The latter ones are considered to be more beneficial for overall learning (Kalamski, 2007; Kinstch, 2004).

Given that the standard of coherence for reading a text is not always optimal (Van den Broek et al., 2001), not all possible inferences will be generated, not even by good readers (Calvo, Estevez, & Dowens, 2003). With the help of reading comprehension assignments, readers can be encouraged to actively process the text so that their standard of coherence will be higher in order to complete the task. These active, higher level tasks help experienced readers generate inferences, remember information and as such, enhance learning from text (Cerdán, Vidan-Abarca, Martínez, Gilabert, & Gil, 2009; Wixon, 1983).

Three often-used reading comprehension tasks that differ with respect to the standard of coherences are gap filling, inference questions, and summarizing. Gap filling involves the completion of randomly deleted words in the text and is highly dependent on surface-level syntactic processes (Carroll, 1972). Inference questions posed after reading the text require more active processing at a local level. And summary writing can be considered a higher order task (Hidi & Anderson, 1986; Westby, Culatta, Lawrence, & Hall-Kenyon, 2010) requiring the integration of large text segments. Summary writing is evidenced to increase both text comprehension and learning in experienced readers (Wittrock & Alesandrini, 1990) and less experienced readers (Franzke, Kintsch, Caccamise, Johnson, & Dooley, 2005).

Reader effects

Two main reader effects that influence learning from text are prior knowledge and working memory. The more prior vocabulary knowledge readers have, the better able they are to extract new word meaning from text (Kintsch, 1994; McNamara, Kintsch, Songer, & Kinstsch, 1996). Previous research has shown that vocabulary is an important predictor in the likelihood and speed of inference generation (Calvo et al., 2003; Singer et al., 1992). Most of these studies were conducted in controlled environments, and often inference generation was prompted by a question. For example, Calvo et al. (2003) found that low vocabulary readers produced less elaborate inferences, and the inferences were generated slower than when readers had more vocabulary knowledge. Singer et al. (1992) found similar results for global bridging inferences, but only a speed effect for local bridging inferences. In addition, several studies have shown that vocabulary knowledge is an important factor in learning from text. Children with low vocabulary knowledge encounter difficulties in acquiring new knowledge from text. In a meta-analysis, Swanborn and De Glopper (1999) showed incidental word learning to be better when readers have larger vocabulary, and this turns out to be the case for both breadth and depth of vocabulary (Vermeer, 2001).

With respect to working memory, it has been found that the better readers are able to keep text representations in memory, the more likely they are capable of learning new word meanings (Daneman & Merikle, 1996; Verhoeven & Perfetti, 2008). The role of working memory is clearly established for inference generation (Cain, Oakhill, Barnes, & Bryant, 2001). Furthermore, there is ample empirical evidence showing that comprehension of children with poor working memory is weakened (Nation, 2007). However, few studies have investigated the role of working memory on *learning* from text. Daneman and Green (1986) found that readers with low working memory had more difficulties in constructing new word meanings. More recently, Cain et al. (2004) found evidence that working memory is also important in word learning for children, especially when global inferences are needed.

To sum up, research has focused on context, tasks *or* reader characteristics. In the present study, an attempt was made to study variation of word learning processes by taking into account all three effect types of and to examine their interaction.

The present study

In the present study, we investigated word learning by comparing learning gains from local and global written contexts. A pretest-posttest design was adapted in which knowledge about target words was tested. Furthermore, we investigated local and global word learning as a function of three often-used reading comprehension tasks which tap into different text representation processes: surface-based gap-filling, text-based inference questions and situation-model based summarizing. A between-subjects approach was adapted in which children performed one of these exercises multiple times. This made sure children were able to adapt the standards of coherence to the task demands. Finally, the influence of vocabulary knowledge and WM was determined.

This study focused not only on the individual contribution of context, task, and reader, but also tested interactions among those variables. Singer et al. (1992) showed a relation between context and reader in a study on local and global inference generation as prompted by a question. To answer the local inference question, the reader needed to generate an inference that connected two adjacent sentences. In the global condition, the two sentences were separated by two intervening sentences. Vocabulary knowledge was an important predictor for both local and global inference generation, whereas WM was only important for generating global inferences. It remains unclear whether similar cognitive abilities underlie incidental word learning from text for less experienced readers. Task and reader have been found to interact as well. Eason, Goldberg, Young, Geist, and Cutting (2012) found that inference making skills contributed to complex questions, while lexical knowledge contributed to less complex questions as well. The present study aimed to extent the investigation of task demands for gap filling, inference questions and summary writing.

There are some indications that context and task are also related. Reading comprehension tasks that focus on different levels of text representation enhance memory at the level involved (Wixon, 1983). Thus, lower-level questions increase the probability that this information is remembered by the reader. Higher-level questions increase the memory of higher level information. More recent work (Cerdán, Vidal-Abarco, Gilabert, Gil, & Rouet, 2008) showed that asking higher-level questions that required integration of information across paragraphs (cf. global contexts), resulted into broader search patterns in the text than lower-level questions that required integration of information within paragraphs (cf. local contexts).

The following research question was addressed in the present study: *In what way are word learning outcomes influenced by context, task and reader effects and in what way do these factors interact?* We formulated three hypotheses with respect to main effects. First, we expected to find words from local contexts to be easier to acquire than words from global contexts (cf. Cain et al., 2004). Following Kintsch's (2004) levels of text representation, we expected deeper processing tasks to be more successful at enhancing word learning from text than more shallow processing tasks, and all tasks to be better than the single reading of a text. Concerning reader characteristics, we hypothesized that children with more vocabulary knowledge learn more new words than children with less vocabulary knowledge. No main effect of WM capacity was predicted, since Singer et al. (1992) only found a contribution for global contexts.

Furthermore, we expected to find interactions between the variables. Children with high vocabulary knowledge were expected to make more inferences about words. In addition, working memory was expected to result into better word learning from global contexts, because children with higher working memory are able to remember larger text segments than children with lower working memory. Children with higher working memory abilities were also expected to benefit more from higher-level tasks, whereas children with lower abilities working memory were expected to benefit more from lower-level tasks. Finally, a relationship was expected between task and context. Higher-level tasks were predicted to contribute to global coherence and hence more words from global contexts are learned compared to low-level tasks.

Method

Participants

Fifth grade classes were recruited by sending letters to 45 Dutch primary schools in the centre of the Netherlands. The five participating schools were first in responding spontaneously to this invitation. Two schools had one fifth grade class and the other three schools all had two fifth grade classes. Some students were excluded from analyses, because they spoke Dutch as a second language (*n* = 3), or because they were diagnosed with dyslexia (*n* = 14) or ADHD (*n* = 1). In addition, participants were deleted due to significantly high WM scores two standard deviations from the mean (*n* = 1) or IQ scores below the 25^{th} percentile (*n* = 2). In total, 149 students (83 girls, 66 boys, M_{age}=10;11 years, age range 10;1-12;2) were included in the analyses. These participants had a normal IQ and scored above the 25^{th} percentile (*Standard Progressive Matrices;* Raven, 1960). All students received a "diploma" for participation.

Materials

Child characteristics

General vocabulary. Vocabulary knowledge was tested by a standardized passive vocabulary knowledge test (*Leeswoordenschattaak*, Verhoeven & Vermeer, 1993). This test consists of fifty multiple choice items in which each word was presented in a short and uninformative context, e.g., 'He sells vegetables'. The students were asked for the meaning of the underlined word. Four multiple choice options were presented including a synonym of the target word, e.g., 'grass', 'green soup', 'salad', and 'edible plants'. Two practice words were discussed prior to the test. Questions regarding the task were answered, though no hints to answers were given. Reported scores are the total number of correct answers.

Working memory (WM). Working memory was measured by a backward digit span memory task (*WISC-III^{NL}*, Kort et al., 2005). The researcher read aloud a string of digits using a falling intonation and pausing one second between the digits. The students were instructed to remember the digits in the reverse order. The strings started short (n = 2) with two attempts for each string length. Whenever children correctly remembered at least one of two strings, the researcher continued with a longer string, adding one digit until a maximum (n = 8) was reached. Each correctly remembered string accounted for one point with a maximum of 14.

5

Task

Reading materials[1]. All participants received six texts (M_{length}= 447 words, *Range* = 374-493) on different diseases adapted from Bouckaert (2007). All texts contained one topic; every text was about a different disease. With the help of a word frequency list (Vermeer, 2000) the texts were manipulated to ensure that all words - except for the target words - were known by typical fifth grade students. The texts were presented with additional line numbers and interlined spacing for comfortable reading. In addition to the experimental texts, a short exercise text was constructed (administered from NCRV, 2008).

From each text, four unknown words were selected: two that could be inferred with help of a local inference, and two that could be inferred with help of a global inference. The words were unknown according to a frequency list representing all words present in school books at primary schools (Vermeer, 2000). All target words occurred

[1] For those interested in the exact items used, all materials are available by e-mail (l.deleeuw@pwo.ru.nl).

at least two times in each text, in both occasions the context was either local or global. For local contexts, the explanation was provided on the first encounter. Words were printed in roman style and hence not explicitly marked.

Subsequent reading comprehension tasks. All participants received several attitude questions on a five-point Likert scale to judge the text on difficulty and likeability. After completing the attitude questions, students - except for the ones in the no task condition - were asked to complete an assignment.

Gap filling task. Students were asked to hand in the text and thereafter received a new copy of the text. In the new copy, every eighteenth word was deleted. Deleted words were listed in random order at the top of the text. The students were instructed to put the words into the right gap in the text.

Question task. The question condition contained three types of questions. Information focus was controlled for by making sure the questions focussed on all text parts and not on target words. First, there were four multiple-choice reference questions with each four answer options. An anaphor with its line number was presented and the children were asked for the proper referent. The second type of questions concerned inference questions. A factual statement was presented to the children and they were asked to confirm or reject it. The students were asked to write down one sentence from the text that helped them answer this question; inferred the answer. The final type of questions was based on sentence integration. One sentence was presented, but the children were told this sentence was a summary of two sentences from the text. It was their job to write down the two sentences that represented the same information that was in the statement.

Summary task. The students were asked to write a summary of the text. Students were instructed that they needed to write a summary containing the most important information. The summary was supposedly helping one of their fellow students who happened to be ill that day. He or she needed to make a test on the text topic and could not read the text, but was allowed to read the summary. The summary was limited to about half a page.

Domain-specific vocabulary knowledge: Target words

To determine vocabulary knowledge of the target words, a vocabulary interview was administered, following the design of Verhoeven and Vermeer (1993). The interview was administered individually and orally in a pretest-posttest design. A list containing the 24 randomly organized target words was constructed. The order of the list remained constant during the experiment to control for order-effects.

For both pretest and posttest, students were asked to tell everything they knew about the word; also small details or minor associations. When the word was unknown, students could tell the researcher that they did not know its meaning. At pre-test, the example word *insect* was used to check the clarity of instruction. When the task was still unclear, the second example, *blind,* was used as an exercise. At post-test no examples were used, but the main purpose was recapitulated. Prior to the target words, the student was explained that the list contained very difficult words they may very well never have heard of.

During the word interview, the researcher did not ask additional questions, except when the student's response was not specific, i.e. 'It is a disease.' In this case, the researcher was allowed to ask if he or she knew what kind of disease it was. Whenever the student presumable did not know anything about the word, i.e., did not respond at all, the researcher specifically asked whether the student had ever heard the word before. If the student responded 'no', the researcher continued the interview. In case of 'yes' the researcher asked whether the child had any associations with the word. All interviews were taped and scored from 0 (*no or false response*) till 4 (*complete understanding of the word*). To determine the reliability of the scoring, one fourth of all interviews were score by two researchers with an intra-class correlation of .93.

Procedure

At the first phase of the study, tests on general vocabulary and WM were administered. Also the domain-specific vocabulary interview was conducted. The first test phase consisted of a written and an oral component, respectively administered in class settings and individually. Prior to the intervention phase, the researcher explained the procedure and the purpose of the tasks using an example text and assignment. The example text was read out loud by one of the students. They were allowed to ask questions. There- after, across a period of two weeks, the students read six texts, equally distributed over the weeks. For every text a similar procedure was adapted. First, the students read the text for at least five minutes. Then, they either received a reading comprehension task (gap filling, questions, or summary) or the session stopped (no task). In the question and summary condition, the students were able to re-read the whole text to complete their assignment. However, in the gap filling condition the teachers collected the original texts before handing out the assignment, as filling in the gaps would not be meaningful when the original text is available. Students could spend a maximum of twenty minutes to complete the assignment. All were informed about the “diploma” they could earn when they completed the assignments. At the final phase, the posttest on domain- specific vocabulary knowledge was administered.

Results

To answer our research question on the effect of context, task and reader on word learning from text, a repeated measures analyses was conducted with Time (pretest, posttest) and Context Type (local, global) as within-subject factor and Task (no task, gap filling, questions, summary) as between-subject factors. Reader differences on vocabulary knowledge and WM were included as covariables and interactions with Time, Context Type and Task were included in the analysis. Four words - two local and two global - were excluded from further analyses, because scores were already high at pretest. Mean scores on general vocabulary and WM are presented in Table 1. Mean scores on pretest and posttest for each task are presented in Table 2.

Table 1

Means and Standard Deviations for General Vocabulary Knowledge and Working Memory

	n	*M*	(*SD*)	*Range*
General vocabulary knowledge (max = 50)	148	38.09	(5.98)	21-49
Working memory (max = 14)	149	5.14	(1.46)	2-9

Table 2

Means and Standard Deviations of Domain-Specific Vocabulary Knowledge on Pretest and Posttest per Task

		Local contexts (max = 4)				Global contexts (max = 4)			
Task		Pretest		Posttest		Pretest		Posttest	
	n	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
No task	35	0.59	(0.42)	0.86	(0.55)	0.85	(0.33)	1.04	(0.42)
Gap filling	39	0.50	(0.30)	0.97	(0.54)	0.76	(0.29)	1.10	(0.49)
Questions	30	0.59	(0.34)	1.16	(0.59)	0.80	(0.36)	1.26	(0.55)
Summary	39	0.56	(0.35)	1.05	(0.57)	0.84	(0.34)	1.28	(0.46)

The results showed main effects of Time, $F(1, 131) = 284.758$, $p < .001$, $\eta^2 = .685$, Context Type, $F(1, 131) = 81.797$, $p < .001$, $\eta^2 = .384$, Vocabulary, $F(1, 131) = 31.967$, $p < .001$, $\eta^2 = .196$, but not for Task, $F(1, 131) = 1.208$, $p = .310$, and WM, $F(1, 131) = .757$, $p = .386$. In addition, interactions were found between Time and Context Type, $F(1, 131) = 4.453$, $p < .05$, $\eta^2 = .033$, and also between Time and Task , $F(1, 131) = 7.528$, $p < .001$, $\eta^2 = .147$. The first interaction indicates that knowledge of local words improved significantly more than the global contexts. Further exploration of the Time X Task interaction with planned contrasts (Dunnett-t) comparing learning gain of each condition to the No Task condition showed that learning gain was higher for the Question ($p < .005$) and Summary ($p < .05$) task than in the No Task condition. Only a marginal effect was found for the Gap Filling task ($p = .068$).

With respect to reader characteristics, we found an interaction between Time and Vocabulary, $F(1, 131) = 24.814$, $p < .001$, $\eta^2 = .159$. Children with higher general vocabulary knowledge were better at learning new words than children with lower vocabulary knowledge. No three-way interactions were found with Time and Context, $F(1, 131) = 1.705$, $p = .194$, or Time and Task, $F(1, 131) = 0.240$, $p < .869$.

Differences in WM capacity showed no interaction between Time and WM, $F(1, 131) = 2.285$, $p = .133$, and no three-way interaction of Time X Context X WM, $F(1, 131) = 1.733$, $p = .190$. However, the interaction of Time X Task X WM was significant, $F(1, 131) = 3.743$, $p < .05$, $\eta^2 = .079$. The observed interaction was further analyzed comparing slope estimates of the interaction learning gain and WM of the tasks against the No Task condition. The outcomes show that for the Question Task, the slope differed significantly from the No Task ($\beta = .510$, $t = 2.235$, $p < .05$), but the slopes of Gap Filling ($\beta = -.119$, $t = -0.540$, $p = .590$) and Summary ($\beta = .170$, $t = 0.734$, $p = .464$) did not. This effect indicates that WM improves word learning in the question condition, but not in the gap filling and summary condition.

5

Discussion

The present study investigated the effects of context, task and reader characteristics on incidental word learning from text. Words inferred from local contexts were compared to words from global contexts. Effects of different processing tasks and individual differences of the reader characteristics were also examined. The results show that words with explanations in local contexts are learned better than words from global contexts. Subsequent higher-order comprehension tasks improved the learning of new words over and above single text reading, but a simple gap-filling task did not add to the learning gain. No differences were found among tasks, and no interaction of task and context was found. With respect to reader characteristics, we found no relation to context. Vocabulary knowledge contributed to overall word learning; with no differences over task. And although WM did not contribute to overall word learning, an interaction was found for the inference question task.

The first hypothesis on main effects was confirmed; learning words from local contexts is less difficult compared to global contexts. This is in line with prior research with experienced readers showing local inference generation is faster than global inference generation (Singer et al., 1992) and word learning is best when the contextual clues are near (Carnine et al., 1984) and context is small (Swanborn & De Glopper, 1999). The results of the present study show that this cost-effect is also pertinent for less experienced readers.

The second hypothesis is partly confirmed: we indeed found that higher-level processing tasks are better for word learning in children than lower-level tasks. However, contrary to our expectation, this lower-level task (gap filling) did not contribute to learning over and above reading the text. Furthermore, the effect of inference questions and summary task did not differ, while we expected summary writing to enhance deeper processing. An explanation for not finding the expected results can be found in task performance. The gap filling task was very easy, with ceiling effects for almost all students. Writing good summaries, on the other hand, might have been too difficult for Grade 5 students, resulting into relatively low quality summaries, and children being less involved in the processing we intended. In order to enhance deeper processing, future studies could include more difficult gap filling tasks, for example by not giving a list of deleted words but leaving the words blank (Pino & Eskenazi, 2009). Furthermore, it can be recommended to train students to write better summaries that enable deeper text processing to improve learning, for example, by providing feedback (Franzke et al., 2005).

The third hypothesis was partly confirmed. We expected higher vocabulary knowledge to benefit word learning and larger WM capacity to enhance word learning from global contexts. Vocabulary knowledge indeed turned out to be important in word learning from text, whereas WM capacity did not influence the amount of words learned; not in local and not in global contexts. Absence of the interaction of WM and context can be caused by the focus on incidental word learning, whereas Singer et al. (1992) studied prompted learning. The processes involved in incidental word learning are different and rely on automatic - and not on prompted - inference generation. The quality of the updated model is important for what type of inference is generated (Van de Broek et al., 2001). Thus, for incidental word learning the standards of coherence seem to be more important than WM capacity.

Finally, an interaction of task and reader was found for WM capacity, but not for vocabulary knowledge. Therefore, the hypothesis that higher vocabulary knowledge and higher WM students benefit more from higher-level tasks was only partially confirmed. The inference question task showed an interaction with WM, indicating that children with more WM capacity are better in learning new words than children with less WM capacity. The task demands might have lead to a cognitive overload for children with low WM capacity (Van Merriënboer & Sweller, 2005). Cognitive load should therefore be considered when designing inference questions. We did not evidence cognitive load problems in the summary task which can be accounted for by the fact that the relationship between situation model processing and working memory capacity is presumed to be relatively weak. Radvansky and Copeland (2001) did not find a relation between WM and situation models and argued this is because situation models are built by sequentially updating the model, a mechanism which does not rely on WM.

A limitation of the present study lies in the measurement of WM capacity by a digit-span task. Chrysobou, Bablekou, and Tsigilis (2011) showed this type of working memory task to explain unique variance on elaborative inference generation after vocabulary was controlled for. However, memory effects for bridging inferences are only found when using a reading span task (cf. Singer et al., 1992). Since the nature of the two memory tasks is different, they might depend on different cognitive capacities. For example, the reading span task includes words and therefore depends also on vocabulary knowledge and syntactic skills. The digit span task relies less heavily on language processing. Future research should include both tasks to determine their contribution to language processing and to determine the cognitive skills that can affect the tasks.

Practical implications of this study are various, although it should be acknowledged that results are limited to one age group and more research is needed to determine extrapolation to other ages. First, this study reconfirms the importance of vocabulary knowledge, as more prior knowledge helps to learn new words. Second, learning new words is best in local contexts. Especially for children with low vocabulary knowledge, working memory might be overloaded in global contexts and word learning is limited. Especially for this group, new words should therefore be offered in local contexts. Lastly, to stimulate word learning it is important to activate deep processing strategies with higher level tasks.

To conclude, this study showed that word learning in fifth grade depends on text, task and reader characteristics. To help children learn new words, explanations should be offered near the word so the meaning can be extracted with the help of a local inference. In addition, children should be encouraged to make connections between sentences or text parts by involving them in higher-level comprehension tasks. It is crucial to help and stimulate children in word learning from text, as vocabulary knowledge lies in the heart of school success.

References

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, *45*, 122-159. doi:10.1080/01638530701792826.

Bouckaert, G. (2007). *Monsterlijke Microben: en hoe ga je ze te lijf?* Uitgeverij Lannoo nv, Tiel.

Cain, K., Lemmon, K., & Oakhill, J. (2004). Individual differences in the inference of word meanings from Context: The Influence of Reading Comprehension, Vocabulary Knowledge, and Memory Capacity. *Journal of Educational Psychology*, *96*(4), 671-681. doi:10.1037/0022-0663.96.4.671.

Cain, K., Oakhill, J., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859. doi: 10.3758/BF03196414.

Calvo, M. G., Estevez, A., & Dowens, M. G. (2003). Time course of elaborative inferences in reading as a function of prior vocabulary knowledge. *Learning and Instruction*, *13*, 611–631. doi: 10.1016/S0959-4752(02)00055-5.

Carnine, D., Kameenui, E. J., & Coyle, G. (1984). Utilization of contextual information in determining the meaning of unfamiliar words. *Reading Research Quarterly*, *19*, 188-203.

Carroll, J. B. (1972). Defining language comprehension: some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language Comprehension and the Acquisition of Knowledge*. New York: Halsted.

Cerdán, R., Vidal-Abarca, E., Gilabert, R., Gil, L., & Rouet, J. (2008). Search and Comprehension Processes in Learning from Text. *Understanding Multimedia Documents*, 121-147. doi: 10.1007/978-0-387-73337-1_7.

Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction*, *19,* 13-27. doi:10.1016/j.learninstruc.2007.12.003.

Chrysobou, E., Bablekou, Z., & Tsigilis, N. (2011).Working memory contributions to reading comprehension components in middle childhood children. *The American Journal of Psychology*, *124*(3), 275-289. doi: 10.5406/amerjpsyc.124.3.0275.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, *25*, 1-18. doi: 10.1016/0749-596X(86)90018-5.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4). 422-433.

Eason, S. H., Goldberg, L.F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, *104*(3), 515-528. doi: 10.1037/a0027182.

Franzke, M., Kintsch, E., Caccamise, D., Johnson, N., & Dooley, S. (2005). Summary street (R): Computer support for comprehension and writing. *Journal of Educational Computing Research*, *33*(1), 53-80. doi: 10.2190/DH8F-QJWM-J457-FQVB.

Fukkink, R. G. (2005). Deriving word meaning from written context: a process analysis. *Learning and Instruction*, *15*, 23-43. doi:10.1016/j.learninstruc.2004.12.002.

Hidi, S., & Anderson, V. (1986). Producing written summaries: Task demands, cognitive operations, and implications for instruction. *Review of Educational Research*, *56*(4), 473-493. doi: 10.2307/1170342.

Kamalski, J. (2007). *Coherence marking, comprehension and persuasion: On the processing and representation of discourse*, Utrecht: LOT.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294-303. doi: 10.1037/0003-066X.49.4.294.

Kintsch, W. (2004). The Construction-Integration model of text comprehension and its implictions for instruction. In R. B. Ruddell, & N. J. Unrau, (Eds.), *Theoretical Models and Processes of Reading*. *Fifth Edition* (pp. 1270-1328). Newark, USA: International Reading Association.

Kort, W., Schittekatte, M., Compaan, E. L., Bosman, M., Bleichrodt, N., Vermeir, G., Resing, W. C. M., & Verhaeghe, P. (2005). *WISC-III NL. Handleiding* [Manual]. *Nederlandse bewerking* [Dutch version]. London: The Psychological Corporation.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*(1), 1-43. doi: 10.1207/s1532690xci1401_1.

McNamara, D. S., & O'Reilly, T. (2009). Theories of comprehension and comprehension skill: Knowledge and strategies versus capacity and suppression. In A. M. Columbus (Ed.), *Progress in Experimental Psychology Research*, Hauppauge, NY: Nova Science Publishers.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research Vol III* (pp. 269-284). New York: Routledge.

Nation, K. (2007). Children's reading comprehension difficulties. In M. J. Snowling & C. Hume (Eds.), *The Science of Reading: a Handbook* (pp. 248-265). Oxford: Blackwell Publishing Ltd.

NCRV (2008, November 29). *Waarom kun je kleurenpotlood niet uitgummen?* [Why can you not remove color pencil?] Retrieved from http://www.willemwever.nl/ vraag_antwoord/ wetenschap-techniek/waarom-kun-je-kleurenpotlood-niet-uitgummen.

Pino, J., & Eskenazi, M. (2009). Measuring hint level in open cloze questions. In H. Chad Lane, & H. W. Guesgen (Eds.), *Proceedings of the Twenty-Second International Florida Artificial Intelligence Research Society Conference, May 19-21, 2009, Sanibel Island,* Florida, USA: AAAI Press.

Radvansky, G. A., & Copeland, D. E. (2001). Working memory and situation model updating. *Memory & Cognition*, *29*(8), 1073-1080. doi: 10.3758/BF03206375.

Raven, J. C. (1960). *Guide to the Standard Progressive Matrices: Sets A, B, C, D and E.* London: H. K. Lewis & Co. Ltd.

Singer, M., Andrusiak, P., Reisdorf , P., & Block, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition*, *20*(5), 539-548.

Swanborn, M. S. L., & De Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review Of Educational Research*, *69*(3), 261-285. doi: 10.3102/00346543069003261.

Tarchi, C. (2010). Reading comprehension of informative texts in secondary school: A focus on direct and indirect effects of reader's prior knowledge. *Learning and Individual Differences*, *20*, 415-420. doi:10.1016/j.lindif.2010.04.002.

Van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition*, *29*(8), 1081-1087. doi: 10.3758/BF03206376.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147-177. doi: 10.1007/s10648-005-3951-0.

Verhoeven, L., & Perfetti, C. (2008). Advances in text comprehension: Model, process and development. *Applied Cognitive Psychology*, *22*, 293-301.doi: 10.1002/acp.1417.

Verhoeven, L., & Vermeer, A. (1993). *Taaltoets Allochtone Kinderen: Handleiding* [*Language Test Immigrant Children: Tutorial*]. Tilburg: Zwijssen.

Vermeer, A. (2000). Lexicale rijkdom, tekstmoeilijkheid en woordenschatgrootte *[Lexical Knowledge, Text Difficulty and Vocabulary Size]. Toegepaste TaalWetenschap in Artikelen*, *64*, 95-105.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, *22*, 217–234. doi: 10.1017/S0142716401002041.

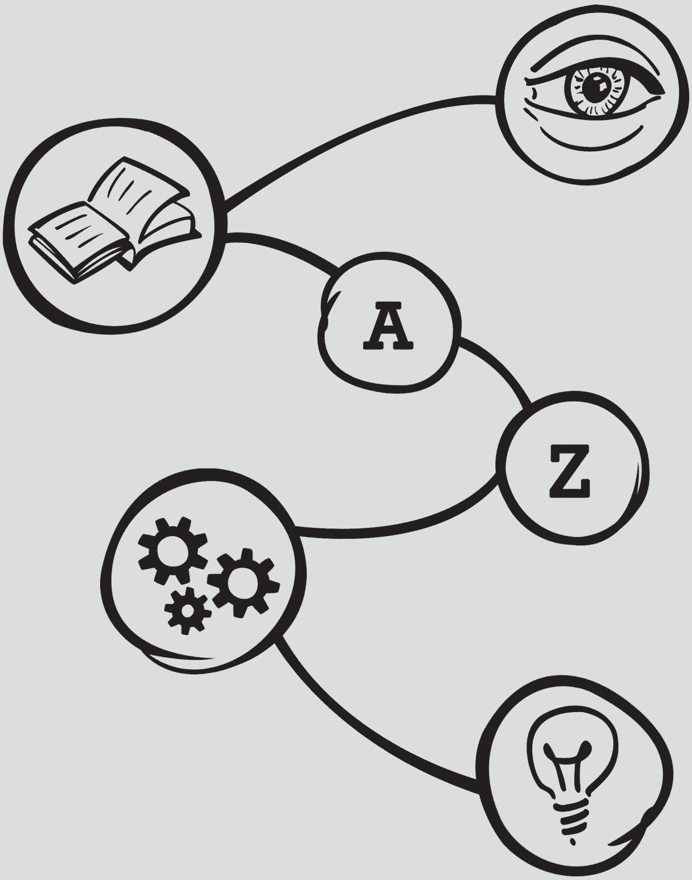
Westby, C., Culatta, B., Lawrence, B., & Hall-Kenyon, K. (2010). Summarizing expository texts. *Topics in Language Disorders*, *30*(4), 275–287. doi: 10.1097/TLD.0b013e3181ff5a88.

Wittrock, M. C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. *American Educational Research Journal*, *27*(3), 489-502. doi: 10.2307/1162933.

Wixon, K. (1983). Postreading question-answer interactions and children's learning from text. *Journal of Educational Psychology*, *30*(3), 413-423. doi: 10.1037/0022-0663.75.3.413

CHAPTER 6

General discussion



General discussion

The main aim of this thesis is to develop further understanding of text comprehension processes by examining how student-, text-, and task-related characteristics influence the process and product of reading. In this final chapter, the role of each of these three characteristics is discussed in light of current theories about reading comprehension. In addition, limitations of the present research and suggestions for future directions are addressed. Finally, practical implications of the present thesis are provided.

Student-related effects on reading

To examine the effects of student-related characteristics on reading processes and outcomes, three eye tracking studies were conducted (Chapters 2, 3, and 4). In previous studies, it has been found that several linguistic abilities (decoding, vocabulary knowledge, and reading comprehension skill) and cognitive abilities (short-term memory, working memory, and non-verbal intelligence) can be considered relevant to reading comprehension processes and outcomes (Cain, Oakhill, Barnes, & Bryant, 2001; Cain, Oakhill, & Bryant, 2004; Daneman & Merikle, 1996).

Reading comprehension models identify *decoding* as an important predictor of reading comprehension (Gough & Tunmer, 1986; Gough, Hoover, Peterson, Cornoldi, & Oakhill, 1996; Perfetti & Stafura, 2014), though the evidence is not very specific concerning whether this skill is important for text processing, learning from texts, or both. Empirical evidence for the importance of decoding for processing is found in all three eye movement studies in the present thesis. Results show that higher decoding skill is related to more skipping and shorter eye movement durations (Chapters 2, 3, and 4). Nonetheless, no direct effect of decoding was found on reading comprehension outcomes. Hence, decoding skill was found to be related to reading comprehension processes, but not to reading comprehension outcomes. With respect to the Simple View of Reading (Hoover & Gough, 1990), it can be concluded that decoding skill itself does not influence reading comprehension outcomes directly, because less skilled decoders are able to compensate their lack of decoding skill by spending additional time reading to end up with a coherent text representation. Hence, decoding efficiency can be associated with the effort a reader needs to put in comprehension, but it does not necessarily predict reading comprehension outcomes (i.e., learning from text).

Vocabulary knowledge can also be seen as an important predictor of reading comprehension. According to the lexical quality hypothesis, it is one of the most essential components (Perfetti & Stafura, 2014). Perfetti and Stafura (2014) state that the quality of word representations (both in linguistic and semantic form) is important for comprehension processes. In our studies, the role of general vocabulary knowledge was found to be related to real-time reading (Chapters 3) and to learning (Chapters 4 and 5). Relatively shorter reading times were found for readers with higher vocabulary knowledge (Chapter 3). Along with these results, it can be concluded that general vocabulary knowledge influences processing and fosters learning from texts, thus supporting the lexical quality hypothesis (Perfetti & Stafura, 2014).

Furthermore, the influences of *reading comprehension* skills on reading comprehension processes and outcomes were examined. Reading comprehension skills can be seen as a collection of skills that are needed to answer reading comprehension questions when the text is present. These skills are found to be very important for reading comprehension outcomes (McNamara & O'Reilly, 2009), although their relation to real-time processes is less clear. The results reported in this thesis indicate that reading comprehension skills are correlated to scores on text comprehension questions. Also, an interaction with word frequency for gaze duration was found in Chapter 3. However, other effects remain insignificant over and above other literacy skills. Therefore, the results of this thesis finds no additional value of reading comprehension skill on comprehension processes and outcomes over and above the several subskills (i.e., decoding, vocabulary).

Previous research is indecisive about the role of memory on reading comprehension, though it is presumed that memory limits the capacity important for comprehension processes (Perfetti & Stafura, 2014). The results reported in this thesis are consistent with respect to *short-term memory*: gaze duration is shorter for students with higher short-term memory capacity (Chapters 2 and 3). However, the results for *working memory* effects were inconclusive. Effects were found for working memory especially in Grade 5, as discussed in Chapter 2; but they were not significant for Grade 4 (Chapter 3) or Grade 6 (Chapter 4). Results on comprehension outcomes showed that memory is important to learning. In Chapter 3, short-term memory was found to predict learning outcomes. In Chapter 5, working memory was found to be important only for learning from answering inference questions. Nevertheless, no such effect was found in Chapter 4. Based on the results of this thesis, it is difficult to draw overall conclusions with respect to the role of memory on reading comprehension.

Text-related effects on reading

To test effects of different text-related characteristics, three experiments were conducted. The study reported in Chapter 2 determined the effect of word class and text difficulty; the study in Chapter 3 examined wrap-up and text-region effects; and Chapter 4 focussed on text-structure and text-length effects. For word class, it is known that adults skip function words more often than content words (Roy-Charland, Saint-Aubin, Klein, & Lawrence, 2007). Chapter 2 confirmed that function words are skipped more often by developing readers, although this was not the case for third graders who read an easy text. Furthermore, regressions were initiated more often at content words, especially in difficult texts. These results validate earlier studies which found that regressions tend to be made when difficulties are encountered in a text (Hyönä, Lorch, & Rinck, 2003).

Text-difficulty effects were found for regression duration measures. This finding suggests that regression durations are longer for more difficult texts. No differences were found at initial processing, which indicates that readers do not necessarily slow down their overall text reading but use regressive behaviour to solve comprehension problems. Overall, this tends to slow down reading processes, but only in the comprehension phase of text processing, not in word-identification processes.

Sentence *wrap-up* effects were found for regression measures in Chapter 3, which shows that integration processes take place at sentence-final words (Kaakinen & Hyönä, 2007). These effects were also related to decoding skills. Students with low decoding skill tend to look back less often; but when they do look back, the regression path duration is longer than that of their more skilled peers. Moreover, as wrap-up effects did not influence text comprehension scores, no conclusions can be drawn with respect to integration effects.

More importantly, the *text structure* was found to be related to reading comprehension. As skilled readers were found to be better at mental model construction (McNamara & O'Reilly, 2009) in the use of reading comprehension strategies (McNamara, Ozuru, Best, & O'Reilly, 2007) and in paying attention to headings (Kaakinen, Hyönä, & Keenan, 2003), it was expected that text structure would influence both real-time processes and comprehension after reading. The results presented in Chapter 3 confirm that real-time processes are influenced by text structure. Relatively longer reading times for headings were found for more skilled readers. Furthermore, whenever a paragraph was nested more deeply, reading times for the heading increased. This indicates that mental model building is more complex for more deeply structured propositions. Nevertheless, this thesis does not provide reasons to conclude that paying more attention to hierarchically more salient headings leads to better comprehension scores.

In agreement with previous research, *text length* does affect real-time processing (Chapter 4). Reading time for the remainder of the paragraph was found to diminish towards the end of the text. Different effects of these diminished reading times were found for fast and slow readers. For fast readers, comprehension scores remained stable across the text; but slow readers decreased comprehension scores at the end of the text. These results suggest that fast readers benefit from building a coherent mental model. It can tentatively be concluded that faster readers benefit from a mental model that is already present (Bell, 2011), whereas slow readers do not, and may experience reader fatigue (Graesser, Singer, & Trabasso, 1994; Van den Broek, Risden, Husebye-Hartman, 1995) or may be involved in mind wandering (Nguyen, Binder, Nemier, & Ardoin, 2014). It could also be the case that slow readers build a very elaborate mental model which enhances the occurrence of mistakes in their mental model or causes memory overload (Van Merriënboer & Sweller, 2005). This might lead to mental model construction problems towards the end of the text.

Task-related effects on reading

A final study (Chapter 5) was conducted to determine the effects of reading comprehension tasks on learning. It is well known that reading comprehension tasks set goals for readers that enhance reading processes by increasing the standard of coherence (Van den Broek et al., 2001). Also, tasks that elicit higher-level reading comprehension processes, such as the situation model (Kintsch, 2004), are found to be better for learning. In Chapter 5, three comprehension tasks were examined that tap into the standards of coherence at different levels: surface code (gap filling task), text-based (inference questions), and the situation model (summary writing) level. The results showed that tasks which address the surface code do not improve reading comprehension over and above single-text reading, whereas higher-order comprehension tasks (i.e., text-based and situation model constructions) improved the learning of new words. These outcomes suggest that reading comprehension tasks which address higher-level processes are more appropriate to enhance learning.

Overall conclusion

The results of the research presented in this thesis support the idea that reading is mainly a bottom-up process. With respect to student characteristics, particularly decoding skill and vocabulary knowledge were found to be important. Text characteristics were also found to influence reading times (word length, word frequency, word type, and text length). Moreover, the sentence wrap-up effect reported in Chapter 3 evidenced bottom-up integration processes (cf., Cognitive-Integration model; Kintsch, 2004), as sentence final words showed increased reading times.

Furthermore, this thesis shows that the processes of younger readers are faster than those of older developing readers. Also, regression probabilities reported in Chapters 2 and 3 were found to increase with grade (3^{rd} = 11.8%, 4^{th} = 15.5%, and 5^{th} = 19.1%). A significant increase was found by comparing 3^{rd} and 5^{th} grade in Chapter 2. Nevertheless, since the present thesis does not include a study that directly compared the probabilities of all grades, it is not possible to apply these results to 4^{th} grade students. This increase in regressions might be due to the fact that older students are more focused on reading for comprehension and are hence better at monitoring their behaviour (McNamara & O'Reilly, 2009). However, as no distinct measures of monitoring behaviour were included in this thesis, no conclusive interpretation of the regression behaviour can be made.

The effects of top-down processes were examined in Chapter 5. This study shows that the alteration of the standard of coherence does affect learning outcomes. However, a well-designed task should encourage students to make inferences at the text-based or situation model level, and not at the surface code level. The results also show that generating text-based inference requires working memory capacity, which makes this task less suitable for students with low working memory. The writing of summaries, on the other hand, was not found to be influenced by student-related characteristics.

Finally, with regards to the *product* of reading, this thesis demonstrates that the *process* of reading predicts scores on comprehension questions over and above student-related capabilities. In Chapters 3 and 4, real-time reading processes were found to moderate the effects of literacy and cognitive capabilities on reading comprehension outcomes. In both Chapters 3 and 4, a moderation of decoding on reading comprehension was evidenced as a function of eye movement measures.

Limitations and implications for future research

The present thesis has several limitations. First, overall conclusions should be considered with care, as different materials and different grade levels were investigated across chapters. It should be acknowledged that the effects of memory and reading comprehension skill were not fully consistent. The contribution of memory may be different across grades, as the results of Chapter 2 suggest; but the possibility that incongruent results are caused by differences in text materials or in experimental setup cannot be ruled out. Conclusions with respect to reading development should also be carefully interpreted, as no longitudinal design was adopted. With an eye to future research, longitudinal experiments would be extremely informative, especially when several reader characteristics are included.

Second, the eye tracker used in these experiments is limited with respect to temporal resolution. By measuring at 120 Hz, detailed information about the reading process may get lost, the temporal sampling error (Andersson, Nyström, & Holmqvist, 2010) is reduced to a similar level as a 1000 Hz eye tracker, taking into account the large number of data points that we have included in the analyses. Nevertheless, it would be useful to confirm our results by using newer equipment that reaches up to 1000 Hz. This would also improve the spatial resolution, which would facilitate the study of within word effects when reading longer texts (in Chapter 4).

Finally, regarding the role of lexical quality, it is important to note that within the present thesis only general vocabulary knowledge and not domain-specific-vocabulary knowledge was measured. The measurement of domain-specific-vocabulary knowledge could be important, as domain-specific knowledge is found to be highly related to learning outcomes (McNamara, Kintsch, Songer, & Kintsch, 1996). This would indicate that the process of reading is influenced by item-specific vocabulary. Still, it is unclear whether these item-specific effects would predict learning over and above general vocabulary knowledge. Therefore, follow-up research should focus on individual differences in both general and specific vocabulary knowledge to gain more insight into their relation to the process and product of reading comprehension.

Implications for educational practice

The present thesis shows that student characteristics influence both the process and the product of reading. Since effects are found of decoding and vocabulary, these skills can be considered to be essential to optimize reading processes. Therefore, training these skills is expected to facilitate easier and faster text processing. In addition, text-related characteristics are found to highly impact the reading process. Difficult texts that contain longer and less frequent words may cause the reader to slow down. Hence, difficult text is not considered optimal for learning; the reader has to put in too much effort to achieve the learning objective. In order to optimize this process, the text should be adapted to the level of the reader.

However, optimizing the reading process does not necessarily result in better text representations or better learning from text. Because decoding and most text-related characteristics were not found to affect reading comprehension outcomes, it seems that faster processing does not always foster comprehension. Nevertheless, vocabulary knowledge and text length were found to be related to reading comprehension scores. Three lessons can be learned. First, the results of this thesis stress that vocabulary training is crucial. Educational practice should therefore focus on vocabulary training as much and as early as possible in order to prevent Matthew effects (Stanovich, 1986). Second, texts should contain sufficient known words but also some new words (Goossens &Vermeer, 2009) to optimize word learning from context. Third, slow readers should be presented with texts of moderate length to prevent overloading when reading for comprehension, because longer texts negatively affect reader comprehension scores in students with low decoding skills.

Lastly, when designing reading comprehension tasks aimed to improve learning outcomes, it is important to encourage students to make inferences across sentences (by answering inference questions or writing a summary, for example). In addition, comprehension tasks should be designed that not only tap into these processes but also take into account the cognitive resources students have available. For example, enhancing inference generation by requesting these inferences very directly might overload the working memory of students with lower working memory capacity. For these students, it might be better to endorse the generation of elaborate inferences in tasks like writing a summary.

Overall, the results of the present thesis suggest that, to improve reading processes, students must be taught decoding skills and texts must be adapted to the level of the children. This helps the reader to optimize his or her reading process, and reduces the effort the reader needs to put into reading for comprehension. However, training decoding or adapting the level does not necessarily improve comprehension. Because the results of this thesis show that vocabulary knowledge is positively related to both eye movements and reading comprehension outcomes, a focus on vocabulary instruction seems to be mandatory to optimize both the process and product of reading comprehension.

References

Andersson, R., Nyström, M., & Holmqvist, K. (2010). Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. *Journal of Eye Movement Research*, *3*(3), 6, 1-12.

Bell, A. (2011). Re-constructing Babel: Discourse analysis, hermeneutics and the Interpretive Arc. *Discourse Studies, 13*(5), 519–568. doi:10.1177/1461445611412699.

Cain, K., Oakhill, J., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859. doi: 10.3758/BF03196414.

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

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4). 422-433. doi:10.3758/s13423-013-0565-2.

Goossens, N., & Vermeer, A. (2009). Wat is een optimale tekstdekking? Woordkennis en tekstbegrip in groep 6 [*What is the optimal text coverage? Vocabulary knowledge and text comprehension in Grade 6*]. *Toegepaste Taalwetenschap in Artikelen*, *82*, 81-92.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*, 6-10. doi: 10.1177/074193258600700104.

Gough, P. B., Hoover, W. A., Peterson, C. L., Cornoldi, C., & Oakhill, J. (1996). Some observations on a simple view of reading. *Reading comprehension difficulties: Processes and intervention*, 1-13.

Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371-395. doi:10.1037/0033-295X.101.3.371.

Hyönä, J., Lorch, R. F., Jr., & Rinck, M. (2003). Eye-movement measures to study global text processing. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye-movement research* (pp. 313-334). Amsterdam: Elsevier Science.

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

Kintsch, W. (2004). The Construction-Integration model of text comprehension and its implications for instruction. In: R.B. Ruddell, & N. J. Unrau, (Eds.), *Theoretical Models and Processes of Reading*. *Fifth Edition* (pp.1270-1328). Newark, USA: International Reading Association.

Kaakinen , J. K. & Hyönä, J. (2007). Perspective effects in repeated reading: An eye movement study. *Memory & Cognition, 35*, 1323-1336. doi: 10.3758/BF03193604.

Kaakinen, J., Hyönä, 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*, 447–457. DOI:10.1037/0278-7393.29.3.447.

McNamara, D. S., & O'Reilly, T. (2009). Theories of comprehension and comprehension skill: Knowledge and strategies versus capacity and suppression. In A.M. Columbus (Ed.), *Progress in Experimental Psychology Research*, Hauppauge, NY: Nova Science Publishers.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are Good Texts Always Better? Interactions of Text Coherence, Background Knowledge, and Levels of Understanding in Learning from Text. *Cognition and Instruction*, *14*(1), 1-43.

McNamara, D., Ozuru, Y., Best, R., & O'Reilly, T. (2007). The 4-pronged comprehension strategy framework. In D.S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 465-496). Mahwah, NJ: Erlbaum.

Nguyen, K., Binder, K., Nemier, C., & Ardoin, S. (2014). Gotcha! Catching Kids During Mindless Reading, *Scientific Studies of Reading*, *18*(4), 274-290. doi:10.1080/10888438.2013.876424

Perfetti, C. A., & Stafura, J. (2014) Word knowledge in a theory of reading comprehension, *Scientific Studies of Reading, 18*, 22–37. doi:10.1080/10888438.2013.827687.

Roy-Charland, A., Saint-Aubin, J., Klein, R. M., & Lawrence, M. (2007). Eye-movements as direct tests of the GO model for the missing-letter effect. *Perception & Psychophysics, 69*(3), 324-337. doi: 10.3758/BF03193753.

Van den Broek, P. W., Risden, K., & Husebye-Hartman E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In: R. F. Lorch Jr. & E. J. O'Brien (Eds.) *Sources of coherence in text comprehension*. Hillsdale, NJ: Erlbaum. 353-373.

Van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition*, *29*(8), 1081-1087. doi: 10.3758/BF03206376.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, *17*(2). doi: 10.1007/s10648-005-3951-0.

Nederlandse Samenvatting

Samenvatting

Begrijpend lezen is een belangrijke vaardigheid, omdat deze vaardigheid helpt bij het begrijpen van informatie uit een geschreven context om deze vervolgens te onthouden. Kunnen lezen met begrip wordt daarom gezien als een belangrijke voorspeller van schoolsucces. Dit komt mede doordat op school veel informatie in teksten wordt aangeboden. Het aanleren van begrijpend leesvaardigheden is daarom op veel basisscholen een speerpunt. Het is daarbij van belang dat kinderen niet alleen leren hoe ze een tekst moeten lezen, maar ook hoe ze de informatie uit de tekst het beste kunnen onthouden. Kinderen moeten dus niet alleen geschreven taal kunnen 'ontcijferen', maar ook snappen op welke manier zinnen en alinea's samenhangen. De samenhang tussen informatie-eenheden wordt opgeslagen in het mentale model dat de lezer van een tekst maakt. Hoe beter het mentale model van de lezer is, hoe waarschijnlijker het is dat de lezer de informatie uit de tekst onthoudt.

Vanaf groep 5 vindt er een omslag plaats van leren om te lezen naar het lezen om te leren. Eerder onderzoek toont aan dat lees- en cognitieve vaardigheden van invloed zijn op het lezen met begrip, waaronder decodeervaardigheden, woordenschat, korte-termijn- en werkgeheugen, begrijpend leesvaardigheden en non-verbaal redeneervermogen. Er is echter weinig bekend over verschillen tussen leerlingen tijdens het begrijpend lezen. Onderzoek bij volwassenen toont aan dat het leesproces van goede en slechte lezers verschilt, zowel op leestempo als op leesstrategie. Ook beïnvloeden tekstkenmerken - zoals tekststructuur en tekstlengte - het leesproces. Tenslotte is het aanbieden van een goede verwerkingstaak belangrijk is om tekstbegrip te bevorderen.

Begrijpend lezen wordt dus beïnvloedt door verschillen tussen lezers, teksten en verwerkingstaken. Het eerste doel van dit proefschrift was om te bepalen hoe het leesproces zich bij beginnende lezers ontwikkelt en welke rol de vaardigheden van de leerlingen hebben. Hierbij werd ook onderzocht in hoeverre het leesproces het begrip van de tekst beïnvloedt. Een tweede doel betrof de invloed van tekstkenmerken op het leesproces en tekstbegrip van beginnende lezers. Hierbij werden verschillende tekstkenmerken getoetst: tekstmoeilijkheid (hoofdstuk 2), tekststructuur (hoofdstuk 3 en 4) en tekstlengte (hoofdstuk 4). Het laatste doel was de effectiviteit van verschillende verwerkingstaken vaststellen (hoofdstuk 5).

Effect van vaardigheden

Het eerste doel van dit proefschrift was te bepalen hoe het leesproces zich ontwikkelt en welke vaardigheden hierbij een rol spelen. Dit is getest in vier experimenten waarbij de vaardigheden van de leerlingen (decodeervaardigheid, woordenschat, begrijpend leesvaardigheid, korte termijn geheugen, werkgeheugen en non-verbaal redeneervermogen) in verband gebracht zijn met het leesproces en scores op tekstbegripvragen. Het leesproces werd in kaart gebracht door de oogbewegingen van leerlingen tijdens het lezen te volgen met behulp van een eye tracker. Tekstbegrip werd gemeten door vragen te stellen over de tekst. In hoofdstuk 2 werden de vaardigheden in verband gebracht met de oogbewegingen van 24 leerlingen uit groep 3 en 20 leerlingen uit groep 5 tijdens het lezen van een gemakkelijkere en een moeilijkere tekst. In hoofdstuk 3 en 4 zijn respectievelijk 40 leerlingen van groep 6 en 73 leerlingen uit groep 8 onderzocht tijdens het lezen van langere informatieve teksten. In het laatste onderzoek (hoofdstuk 5) werd bij 149 leerlingen van groep 7 het effect van vaardigheden in een interventiestudie bekeken.

In alle experimenten gepresenteerd in dit proefschrift is evidentie gevonden voor het feit dat leesprocessen gerelateerd zijn aan decodeervaardigheden; betere decodeerders hadden kortere leestijden. Deze leerlingen waren overigens niet alleen sneller, maar sloegen ook vaker woorden over (hoofdstuk 2). Echter, in geen van de onderzoeken werden aanwijzingen gevonden dat dit snellere leesproces leidt tot beter begrip. Kortom, we kunnen zeggen dat goede decodeervaardigheden helpen bij het verhogen van het leestempo, maar het lijkt er niet op dat dit proces ook efficiënter is dan het leesproces van langzamere lezers.

Verder tonen de resultaten uit dit proefschrift aan dat woordenschat een belangrijke vaardigheid is. Voor leesprocessen werd een effect gerapporteerd in hoofdstuk 3, wat laat zien dat een goede woordenschat verband houdt met snellere leestijden. Deze effecten vonden we overigens niet in hoofdstuk 2 en 4. Er zal dus er meer onderzoek gedaan moeten worden gedaan om te bekijken op welke manier woordenschat het leesproces beïnvloedt. Duidelijkere effecten van woordenschat werden gevonden in relatie tot tekstbegrip; woordenschat droeg in alle gevallen bij aan het voorspellen van tekstbegrip.

Het effect van begrijpend leesvaardigheden werd ook getoetst in dit proefschrift. Belangrijk hierbij is dat begrijpend leesvaardigheden veelal een verzameling zijn van een aantal deelvaardigheden (decoderen, woordenschat, enzovoorts). Het speciale aan begrijpend leesvaardigheden is echter dat deze vaardigheden gecombineerd moeten worden ingezet tijdens het lezen. Dit proefschrift laat geen toegevoegde waarde zien van die gecombineerde vaardigheid; er werd geen verband gevonden tussen begrijpend leesvaardigheden en het algemene leesproces. Goede lezers lijken dus niet per definitie sneller te lezen. Wel lieten de studies in hoofdstuk 3 en 4 zien dat leerlingen met betere begrijpend leesvaardigheden hun aandacht anders verdelen over de tekst. Ze besteedden meer aandacht aan belangrijke tekstdelen, zoals de titel van de tekst, in vergelijking met kinderen die slechtere begrijpend leesvaardigheden hebben. Ook werd er geen effect gevonden van begrijpend leesvaardigheden op tekstbegrip. Uit de resultaten kan dus afgeleid worden dat begrijpend leesvaardigheden geen toegevoegde voorspellende waarde hebben bovenop de andere deelvaardigheden.

Conclusies over de rol van het korte-termijn en werkgeheugen in begrijpend leesprocessen zijn lastig te bepalen op basis van dit proefschrift. Aan de ene kant werden er effecten van zowel korte-termijn geheugen en werkgeheugen op het leesproces gevonden in hoofdstuk 2, en op tekstbegrip in hoofdstuk 3. Aan de andere kant bleven deze effecten uit in de andere hoofdstukken. Een mogelijke verklaringen hiervoor kan zijn dat het effect van geheugen afhankelijk is van de leeftijd van de leerlingen. Deze verklaring is deels in overeenstemming met de resultaten in hoofdstuk 2, waar werd aangetoond dat werkgeheugen een grotere rol speelt bij leerlingen uit groep 7 dan uit groep 5. Een andere verklaring kan zijn dat geheugeneffecten sterk afhankelijk zijn van de context. Dit is in overeenstemming met de inconsistentie van dit type effecten in eerder onderzoek; soms worden er wel en soms geen effecten van geheugen op begrijpend lezen gevonden. Dit proefschrift sluit echter geen van deze verklaringen uit.

Als laatste is er gekeken naar de invloed van non-verbaal redeneervermogen op begrijpend leesprocessen en tekstbegrip. In dit proefschrift zijn er geen aanwijzingen gevonden dat redeneervermogen invloed heeft op het leesproces. Wel werd er een effect gevonden op het tekstbegrip; kinderen met een laag redeneervermogen beantwoordden de vragen van een tekst slechter wanneer ze meer woorden oversloegen. Het lijkt voor deze groep dus vooral belangrijk dat ze niet proberen de tekst zo snel mogelijk, maar juist zo secuur mogelijk te lezen.

Effect van tekstkenmerken

Een tweede doel van dit proefschrift was het bepalen wat de invloed van de tekst op het leesproces. In ieder hoofdstuk stond een ander tekstkenmerk centraal. In hoofdstuk 2 werd het lezen van een gemakkelijke en een moeilijkere tekst met elkaar vergeleken. Eerder onderzoek toonde reeds aan dat het leesgedrag verandert wanneer een lezer wordt geconfronteerd met een moeilijk stuk tekst en de lezer zal zijn of haar tempo naar beneden aanpassen. Ook zal de lezer in een moeilijke tekst vaker problemen hebben met tekstbegrip, met als gevolg dat de lezer vaker zal terugkijken in de tekst. Uit de resultaten van hoofdstuk 2 bleek inderdaad dat het lezen van een moeilijkere tekst anders verloopt dan het lezen van een gemakkelijkere tekst. Het verschil werd vooral zichtbaar in het terugleespatroon; leerlingen gingen vaker terug om tekstdelen opnieuw te lezen. Ook was de tijd die ze besteedden aan het teruglezen langer voor de moeilijkere tekst. Dit experiment toont daarmee aan dat leespatroon van kinderen afhankelijk is van de vaardigheden van de lezer, maar ook van de tekst zelf.

In hoofdstuk 3 werd een volgend tekstkenmerken onderzocht: tekststructuur. Er werd gekeken naar effecten binnen zinnen en binnen alinea’s. Bij de effecten binnen zinnen werden woorden die aan het eind staan vergeleken met de andere woorden. Eerder onderzoek toont aan dat volwassen lezers langzamer gaan lezen wanneer ze aan het eind van de zin komen. Dit doen lezers, omdat ze aan het einde van een zin de informatie ervan zullen integreren in hun mentale model. De resultaten zoals gepresenteerd in hoofdstuk 3 lieten zien dat er bij woorden aan het einde van de zin vaker en langer werd teruggelezen, wat evidentie is voor het integratieproces aan het einde van een zin. Daarnaast was ook gekeken naar effecten van segmenten binnen alinea's. De leestijden van de titel, de eerste zin van de alinea, en de laatste zin van de alinea werden vergeleken met de tussenliggende zinnen. De resultaten lieten zien dat titelwoorden minder vaak werden overgeslagen en dat er langer wordt gelezen. Dit effect was groter bij goede lezers. Ook wordt er niet vaker, maar wel langer teruggelezen vanaf de laatste zin van de alinea. Dit effect was groter bij leerlingen met lage decodeervaardigheden. Opvallend was dat er geen effecten van begrijpend leesvaardigheden op tekstbegrip werden gevonden. Dus ondanks dat het leesproces anders is voor verschillende tekstsegmenten, beïnvloedt dit niet de scores op begripsvragen na het lezen van de tekst.

In hoofdstuk 4 werd een onderzoek beschreven waarin leerlingen uit groep 8 informatieve teksten lazen bestaande uit tien alinea’s. De tekst begon met één introducerende alinea. Daarna werden de drie thema’s behandeld. Elk thema bevatte drie alinea’s: in de eerste werd het thema geïntroduceerd en in de twee volgende alinea’s werden twee subthema’s omschreven. Elke alinea was voorzien van een titel. Er werd gekeken of het lezen van de titel en de rest van de alinea veranderde gedurende de tekst. Dit was inderdaad het geval. Zo was de leestijd van de titel afhankelijk van de structuur van de tekst; titels van dieper gestructureerde alinea’s lieten langere leestijden zien. Ten tweede gingen leerlingen steeds minder tijd besteden aan het lezen van de rest van de alinea naarmate ze dichter bij het einde van de tekst kwamen. Dit leidde bij langzame lezers tot slechtere scores op de begripsvragen, terwijl er bij snelle lezers geen verschil was tussen vragen die het begin en einde betroffen. Het leesproces versnelt dus naarmate de tekst langer wordt, maar dit heeft een negatief effect op tekstbegrip voor relatief langzame lezers.

Effect van verwerkingstaken

Het laatste doel richtte zich op de vraag welke opdracht leerlingen uit groep 7 het beste helpt om informatie uit een tekst te verwerken. Vier groepen werden met elkaar vergeleken. De eerste drie groepen maakten bij iedere tekst een opdracht, elk gericht op een bepaald niveau van tekstverwerken: een gatentekst (zinsniveau), het maken van inferentievragen (alineaniveau) en het schrijven van een samenvatting (tekstniveau). De laatste groep maakte geen taak na het lezen van de tekst. Uit de resultaten bleek dat het maken van een opdracht het leren van nieuwe woorden bevordert. Dit was echter alleen het geval voor de inferentievragen en het schrijven van de samenvatting en niet voor de gatentekst. Hieruit blijkt dat opdrachten de lezer moeten stimuleren om ten minste verbanden te leggen op alinea of tekstniveau.

Conclusies en implicaties voor de onderwijspraktijk

De studies in dit proefschrift tonen aan dat individuele verschillen tussen leerlingen invloed hebben op processen tijdens het lezen en resultaten op begripstaken na het lezen. Hierbij wordt leesproces wordt vooral gestuurd door decodeervaardigheden en tekstbegrip vooral door woordenschat. Om het leesproces te versnellen kan dus het beste getraind worden op decodeervaardigheden. Belangrijk is echter dat het leesproces geen directe invloed lijkt te hebben op het leerresultaat. Het stimuleren van sneller lezen leidt dus niet tot beter begrip. Sterker nog, wanneer leerlingen met slechte decodeervaardigheden sneller gaan lezen, kan dit zelfs leiden tot een slechter tekstbegrip. Leerkrachten wordt aangeraden om juist deze groep goed in de gaten te houden tijdens het lezen. Door deze leerlingen rustiger te laten lezen, zullen zij de tekst beter begrijpen.

Om tekstbegrip te bevorderen zal er getraind moeten worden op woordenschat. Het hebben van een hogere woordenschat helpt bovendien bij het sneller lezen. Een goede manier om nieuwe woorden te leren is het aanbieden van nieuwe woorden in een geschreven context. Het is daarbij wel van belang dat de tekst wordt ondersteund met een goede begrijpend leestaak. Een goede begrijpend leestaak richt zich op het leggen van verbanden op hogere niveaus; verbanden tussen zinnen of tussen alinea's. Zo worden leerlingen gestimuleerd om nieuwe woorden uit de tekst te onthouden, waardoor hun woordenschat verder groeit.

Leerkrachten wordt aanbevolen om naast het trainen van vaardigheden de tekst zoveel mogelijk aan te passen op het niveau van de leerlingen. De tekst kan het leesproces namelijk verslechteren. Zo leiden moeilijke teksten tot langere leestijden. Lange teksten kunnen bij langzame lezers ook leiden tot slechter begrip. Het advies is dan ook om leerlingen die moeite hebben met lezen niet te lang achter elkaar te laten lezen, want dit heeft een averechts effect. Misschien is het goed om deze leerlingen een gemakkelijkere tekst te laten lezen of tussendoor een korte opdracht te laten maken om het leesproces en tekstbegrip te bevorderen.

Dankwoord

Dankwoord

Dit proefschrift was er niet gekomen zonder de hulp van een heel veel lieve mensen. Geweldige mensen hebben mij de afgelopen jaren gevraagd (en ongevraagd) voorzien van adviezen, hebben mij bewust, of onbewust, gemotiveerd, geholpen of hebben gezorgd voor de nodige afleiding. Ik ben blij dat ik in dit dankwoord de mogelijkheid heb een aantal van hen te bedanken.

Allereerst wil ik graag mijn begeleiders bedanken. **Ludo**, jij bewaakte altijd de grote lijnen, was de drijfveer in het opzetten van nieuwe designs en bracht mijn resultaten in verband met verschillende wetenschappelijke theorieën. Jij gaf me het vertrouwen dat het goed zou komen. Jouw positiviteit werkt aanstekelijk en gaf me altijd weer de moed om verder te gaan. **Eliane**, bedankt voor alle feedback. In het begin schrok ik soms nog wel eens van de hoeveelheid feedback, maar ik ging het steeds meer waarderen. En omdat jij bruggen bouwde op plekken waar ik te impliciet was, is mijn proefschrift ook voor anderen coherent en leesbaar geworden!

Graag wil ik de manuscriptcommissie bedanken voor het lezen van mijn proefschrift. Met dank aan **Peter-Arno**, **Ted** en **Menno**, heb ik de laatste puntje nog op de i kunnen zetten voordat het proefschrift naar de drukker is gegaan.

Alle scholen die hebben meegewerkt aan mijn onderzoek verdienen ook zeker een plaats in mijn dankwoord. Zonder de toewijding van de schoolleiding, docenten en natuurlijk de leerlingen was het niet mogelijk geweest om dit proefschrift te maken. In totaal hebben maar liefst 13 scholen en bijna 400 leerlingen aan mijn onderzoek deelgenomen. En zonder de hulp van **Lisette**, **Suzan**, **Rozemarijn**, **Lucy**, **Freya**, **Esther** en **Lonneke**, was het me niet gelukt om alle data te verzamelen, te verwerken en te presenteren in dit proefschrift.

Verder heb ik ook veel steun gehad van collega's. Mijn Radboud collega's hielpen me niet alleen met praktische en theoretische vraagstukken. Ze hebben er ook voor gezorgd dat mijn promotietijd heel leerzaam, maar ook vooral heel gezellig, leuk, grappig en inspirerend was. Super fijn dat je gewoon even binnen kan lopen voor een gezellig praatje over van alles en nog wat (**Carmen**, **Brigitte**, en **Elise**). Ook zal ik de samenwerking met **Marco** en **Marjolijn** missen. Met jullie is het altijd lachen! De schrijfweken zijn ook memorabel. Tijdens de eerste schrijfweek (**Eva**, **Loes**, **Gesa**, **Suzan**, **Kim**, **Mark**, **Thys** en **Sabine**), en tijdens het 'knallen in Halle' met **Frauke**, **Evelien**, **Nicole**, en **Karly**. Jammer dat we elkaar toen pas echt goed leerden kennen, want in die week zijn hele fijne vriendschappen ontstaan (met dank aan Sneeuwwitje en

Party & Co). Ook mijn fijne Malmberg collega's hebben me geholpen bij de laatste loodjes van dit proefschrift. **Mark**, **Rianne**, **Thomas**, **Mark**, en **Floor** (maar ook de andere Malmbergers) hebben me vanaf de eerste dag opgenomen in hun team. Ik voel me echt thuis bij VO en ik hoop met mijn kennis en vaardigheden ervoor te zorgen dat Malmberg digitaal marktleider wordt en blijft!

Mijn vrienden zorgden gelukkig voor de nodige afleiding. **Jaap**, die altijd in was voor een glas wijn en een goed gesprek, maar het ook aandurfde om de voorkant van dit proefschrift te ontwerpen! Je bent een held! Maar ook mijn BFF **Merel**, **Rogier**, **Dik**, **Jeroen**, **Bart**, **Joleen**, **Angela**, **Tim**, **Olaf** en **Ilse** hebben mij afgeleid met lekkere etentjes of heerlijke vakanties. Super gezellig :) Leerzaam (én gezellig) was Rijnhuyse. Het organiseren van het jeugdtoernooi (**Annemieke**, **Raymond**, **Christine**, **Annemiek** en **Erwin**) en het opzetten van de nieuwe communicatiemiddelen samen met **Job**, **Mike** en **Paula** hebben mij vooral veel voldoening en inspiratie gegeven.

Mijn familie stond ook altijd voor mij klaar. Zo leende **Marjolijn** mij een fiks bedrag zodat ik een auto(otje) kon kopen om naar Nijmegen op en weer te reizen. En tijdens de vakanties van **Owen** en **Sandra** mocht hun lease-auto lenen, wat vooral fijn was vanwege de tankpas;). Trots ben ik op **Jeroen** en **Shereen**. Jullie hebben laten zien dat jullie echt doorzetters zijn. Kanjers! Op stressvolle moment kon ik altijd vertrouwen op **Anton**. Hij legde mij dan extra in de watten door mijn lievelingseten te koken. **Papa**, bedankt voor je hulp bij het maken van de AWK-scriptjes! En ook **Marianne** en **Gé** hebben altijd in mij geloofd en hebben mij vanaf het eerste begin gesteund.

In de laatste fase van mijn project, bij het aanvragen van een take-off subsidie om mijn eye track kennis in de praktijk te brengen, heb ik hulp gehad uit verschillende hoeken. **Rick** adviseerde mij op het ondernemersvlak en gaf mij extra vertrouwen in mijn idee. **Maarten** (en **Marjolijn**) hielpen mij met het technisch uitdenken van het plan en **Erik** bouwde voor mij een prototype. Als kers op de taart hielp **Josefien** mij met het maken van een promofilmpje waarmee we subsidie definitief binnenhaalden! En tijdens de uitvoering kon ik niet zonder de steun van **Willemijn**. Heel leuk om zulke enthousiaste mensen om mij heen te hebben.

In het bijzonder wil ik graag mijn paranimfen bedanken. Jullie hebben mij door de moeilijkste momenten heen gesleept. **Esther**, met jou heb ik heel wat thee gedronken, problemen besproken en gefilosofeerd over theorieën en nieuwe onderzoeksopzetten. Ook heel erg bedankt dat je mij betrokken hebt (en hopelijk ook zal blijven doen) bij Letterprins! En **Roos**, bij jou kon ik altijd terecht met moeilijke kwesties en jij kan mij als geen ander een spiegel voorhouden. Hierdoor kon ik niet alleen beter problemen oplossen (zoals de marktplaats-oplossing), maar heb ik ook veel geleerd over mezelf! Ik hoop dat je dit wilt blijven doen en dat we nog veel fijne etentjes en vakanties met elkaar zullen hebben.

Mama, je bent de sterkste vrouw die ik ken en je hebt mij altijd de kans gegeven om mijzelf te ontplooien, ook al moest je je daarvoor soms in verschillende bochten wringen. Jij hebt mij laten zien hoe je zelfstandig moet zijn, hoe je door moet zetten als het even lastig is, maar vooral hoe je met hard werken en doorzettingsvermogen tot meer in staat bent dan je denkt. Jij bent daarin altijd mijn grootste voorbeeld geweest!

Lieve **Rutger**, ondanks dat je graag als eerste genoemd had willen worden, sta je nu als allerlaatste, maar wel als allerbelangrijkste in dit dankwoord. Jij bent degene waardoor dit proefschrift uiteindelijk af is gekomen; doordat je er altijd voor me bent, achter me staat, aanmoedigt, meedenkt, doorslaapt als ik 's nachts doorwerkte, maar me ook terugfluit als ik te veel hooi op m'n vork neem. Ik kan met jou alles delen, boos worden, verdrietig zijn, leuke dingen doen en genieten van het leven. Maar vooral wil ik je bedanken omdat je gewoon het liefste vriendje voor me bent!

Curriculum Vitae

Curriculum Vitae

Linda de Leeuw is geboren op 30 september 1985 in Nieuwegein. Na het behalen van het tweetalig VWO op het Anna van Rijn College in Nieuwegein, studeerde zij Communicatie- en Informatiewetenschappen aan de Universiteit Utrecht waarin ze een major Communicatie en een minor Taalkunde volgde. Aansluitend ronde ze in Utrecht de Research Master Linguistics af. Haar scriptieonderzoek, onder begeleiding van Ted Sanders en Pim Mak, betrof de invloed van impliciete causaliteit op het verwerken van pronomina in causale relaties. In september 2009 startte ze met haar promotieproject aan het Behavioural Science Instituut van de Radboud Universiteit te Nijmegen. Tijdens haar project werd ze begeleid door Eliane Segers en Ludo Verhoeven.

Naast haar promotieonderzoek was Linda ook docent bij Pedagogische Wetenschappen. Naast het vak Academische Vaardigheden begeleidde ze verschillende scriptiestudenten van zowel de Bachelor als de Master Pedagogische Wetenschappen. Bovendien heeft zij vanaf 2013 ook meegeholpen met de ontwikkeling en toetsing van de leesapp Letterprins, welke in datzelfde jaar de Nationale Alfabetiseringsprijs won.

Momenteel werkt Linda samen met Ludo Verhoeven, Marjolijn van Hulzen en het Expertisecentrum Nederlands aan de ontwikkeling van de Leesscan. De Leesscan is een programma waarmee spellende en radenende lezers op basis van hun oogbewegingspatroon kunnen worden gesignaleerd. Daarnaast is ze als business consultant werkzaam bij Uitgeverij Malmberg, waar ze advies geeft over het ontwikkeling en toetsing van digitale leermiddelen in het voortgezet onderwijs.

Radboud University

IN·DEI·NOMINE·FELICITER

Radboud University