Reading comprehension in elementary school children: cognitive studies of the reader, the text, and the task

Karlsson, A.K.J.

Author: Karlsson, J.A.K.
Title: Reading comprehension in elementary school children: cognitive studies of the reader, the text, and the task
Issue Date: 2021-04-07

Reading comprehension in elementary school children:

Cognitive studies of the reader, the text and the task

Josefine Karlsson

Title: Reading comprehension in elementary school children:
Cognitive studies of the reader, the text, and the task

Print: Örebro University/Repro 03/2021

Cover illustration: Anna Billing and Colin Demain
Cover models: Felix and Mons

Reading comprehension in elementary school children:

Cognitive studies of the reader, the text, and the task

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 7 april, 2021
klokke 11.15 uur
door
Josefine Anna Karin Karlsson
geboren te Ryd, Zweden
in 1984

Promotor:
Prof. Dr. Paul van den Broek
Co-promotor:
Dr. Linda van Leijenhorst

Promotiecommissie:
Prof. Dr. Eddie Denessen (Universiteit Leiden)
Dr. Jacqueline Evers-Vermeul (Universiteit Utrecht)
Dr. Marga Sikkema-de Jong (Universiteit Leiden)
Prof. Dr. Eliane Segers (Radboud Universiteit)
Prof. Dr. Judith Rispens (Universiteit van Amsterdam)
Dr. Jochanan Verbeek (Universiteit Leiden)

Table of Contents

Chapter 1

General Introduction

1.1 General Introduction

Reading comprehension is a multifaceted skillset important to acquire in order to participate in modern society; to learn at school, for work related communication, for social digitized interactions, and to keep up to date with news. Important developmental change in this skillset occurs between the ages of 9 and 12, when elementary school children go from learning to read to reading to learn. In this phase educators start expecting the children to use their reading comprehension skillset to gather knowledge about many different topics. This knowledge helps children in understanding their current surroundings as well as prepares them for future possibilities of employment. However, children are of course not alone on their journey to become proficient readers. A great deal of research and educational resources are mobilized to help them on their way. With this doctoral dissertation I aim to enlarge the scientific knowledge of reading comprehension and aid educational practitioners who ground their pedagogical work in scientific literature. This gathering of four empirical papers presents research from a cognitive scientific perspective on three elements that are important to understand reading comprehension in scientific and educational contexts: the reader, the text, and the task (Snow & RAND, 2002; van den Broek, Fletcher, & Risden, 1993).

Within cognitive science, gaining deep comprehension of a text is described as the construction of a mental model, a situation model (Johnson-Laird, 1983; van Dijk & Kintsch, 1983). This means that the situations, events, and characters that are depicted in the text need to be envisioned in the reader's mind. Being able to construct a situation model depends on reader characteristics, text characteristics, and task demands. Many cognitive skills and strategies are needed for a reader to construct a situation model. An example of a crucial skill needed to construct a situation model is the ability to make inferences from the text (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Graesser, Kintsch, 1994; Singer, & Trabasso, 1994). Inference generation entails reading between the lines. This requires the reader to connect different parts of the text to other parts of the text, and to connect parts of the text to their previous knowledge. By generating inferences, i.e. seeing how sentences are interconnected and enriching the situation model with previous knowledge, the reader understands that the text is not just a string of words, but that it tells a story with evolving events and with causes and effects. In addition, differences in text topic, difficulty, and length have an impact on what a reader can extract from the text, and thereby how rich the situation model becomes that they are constructing (e.g. van den Broek et al., 1993). In reading an easy text, for example a text that uses everyday language and has clear structure, the reader can easily construct a rich situation model. Thereby, the reader achieves a good understanding of the content. However, as soon as the text becomes more demanding, perhaps introducing new words and concepts, the reader needs to work harder to achieve a good understanding of the text. Finally, the task that the reader is given, or takes on, while reading will have an impact on the ability to construct a rich situation model (e.g. van den Broek & Kendeou, 2017). Reading a text at own free pace and being able to revisit sentences that are hard to understand, will rule a different outcome in comprehension than when trying to understand the same text but without the ability to control the speed with which the text is presented. For example, when being read to or when using a digital read-aloud device. The many complex cognitive processes that contribute to the development of reading comprehension, such as the ability to make inferences during reading, likely interact with reader, text, and task characteristics (Rapp & van den Broek, 2005).

A fair amount of scientific knowledge on the interaction of reader characteristics, text characteristics, and task demands has been gathered studying adult readers, often university students who are quite proficient readers. The research in this dissertation builds on the existing research by predominantly focusing on 9-12-year-old children's abilities to construct a situation model of text. This is an important phase in a child's reading development. 9-12-year olds in many western educational systems, as in the Dutch school system, transition from learning to read to reading to learn (e.g. Chall 1983; 1996; Poolman, Leseman, Doornenbal, & Minnaert, 2017). In the first phase, much of reading instructions focus on the relationship between phonemes and graphemes to decode letters and words; this enables the reader to decipher what the written signs in the text represent. In this first phase of learning to read, children often encounter simple texts. In the second phase, reading to learn, children are expected to have become fluent in word-decoding skills, and are given texts with the aim to teach content. However, these texts also require additional reading processes compared to the simple texts used in the first phase. Reading tasks that the children encounter in upper elementary school lay the knowledge foundation that is necessary for profession-oriented education later in life. With careful consideration of how reading instructions are constructed by educational practitioners, children can learn about many topics and start building knowledge in fields of their interest. The more proficient readers feel, the more motivated they are to keep reading and gather knowledge in fields of their interest (e.g. Willingham, 2015). However, a phenomenon called the fourth-grade slump, in which children's reading performance drops, indicates that the increased demands that come with reading more complicated texts is not gradual (Chall 1983; 1996). Although children at upper elementary school may experience increased demands and opportunities driven by school and home environment, their cognitive system is still developing. Executive functions, and brain networks supporting language skills and executive functions, continue to develop in older children and well into late adolescence (e.g. Diamond, 2013; Gathercole, Pickering, Ambridge, Wearing, 2004; Huizinga, Dolan, & van der Molen, 2006; Zielinski, Gennatas, Zhou, & Seeley, 2010). Executive functions can be described as an umbrella term for cognitive processes that allow control of thoughts and behavior (Diamond, 2013). The main executive functions are working memory, cognitive flexibility, and inhibition. Working memory can be seen as a mental workspace that enables keeping information in mind, defined by a certain storage capacity, and manipulating this information by processes such as updating the content in working memory (Daneman & Merikle, 1996; Diamond, 2013). Behavioral studies of working memory show that its development continues throughout childhood and adolescence (e.g., Diamond, 2013; Gathercole, 1999; Huizinga et al., 2006). Working memory is important in the context of reading comprehension as it underlies the capacity and processes necessary to retain and update the content of the story as the reader proceeds through the text (e.g., Carretti, Borella, Cornoldi, & De Beni, 2009; Daneman & Merikle, 1996). For educational institutes to enable school children to become proficient readers and thereby take part in society, it is important that science provide answers to which processes and strategies children at this age use successfully and unsuccessfully during reading, and under which circumstances (Snow & RAND, 2002). Pinpointing differences between readers in applying such strategies and processes helps understanding why some readers succeed and some struggle with comprehending text. Together, the four papers in this dissertation aim to help in understanding the effect of reader characteristics, text characteristics, and of task demands on reading comprehension in upper elementary school children.

1.2 The Reader

Readers engage in various cognitive processes during reading. Cognitive reading research distinguishes between online processes during reading and the resulting offline situation model. Online processes define what the reader does during reading, whereas offline comprehension defines the gathered understanding that the reader has built up when having finished reading. In this section I describe online processes that readers engage in to build a situation model, and why they are important to study in a developmental population. Although models on reading comprehension differ in the specifics of the relationship between online processes and the offline situation model, a consensus has emerged that online reading processes affect the reader's offline situation model (McNamara & Magliano, 2009). In addition, readers differ in their use of online processes (voluntarily or automatically), which means they may end up with differences in their mental model of the text

(McNamara & Magliano, 2009). Various online processes are needed for a reader to construct a situation model to understand the text. In cognitive science, a distinction is made between lower-order and higher-order online cognitive processes. According to the simple view of reading (e.g. Hoover & Gough, 1990), word decoding entails lower-order processes such as orthographic and phonological awareness. These processes are crucial to transform the written letters into speech sounds and hence decipher the words in a text. Higher-order processes are necessary for the reader to understand the meaning of the text as a whole. An example of a higher-order process is inference generation: readers make backward inferences, i.e. connecting the focal sentence with previous text, and forward inferences, i.e. predicting upcoming content of the text (e.g. van den Broek et al., 1993). Further examples of higher-order processes are keeping track of temporal, spatial and causal dimensions of the text, and updating our understanding of the text (Zwaan, Langston, & Graesser, 1995). Whereas lower-order processes often become automatic with sufficient practice, higher-order processes require a higher degree of executive control, making demands on for example a readers' working memory (Daneman & Merikle, 1996; Daneman & Carpenter, 1980; McNamara & Magliano, 2009). Therefore, it is important to research the relation between reading and working memory. Especially in an age group in which working memory is still developing. In this dissertation, the focus is on studying higher-order cognitive processes, in part in relation to working memory, for an enhanced scientific understanding of comprehension processes in developing readers.

1.2.1 Generating Inferences When Building a Situation Model

The ability to generate the correct inferences during reading is crucial for successful text comprehension (for an anthology on the topic see O'Brien, Cook, & Lorch, 2015). Written (and spoken) communication often assumes that the receiver of the message has some previous or contextual knowledge on the topic. A text usually contains conceptual gaps with implicit meaning. Therefore, to understand a text fully, readers need to fill in these gaps by generating inferences. When making inferences, readers can use previous text, previous knowledge, and text structure knowledge to understand the focal sentence and to make predictions of the text to come. Cognitive reading models describe that from the moment the reader encounters the first sentence in a text, there is a spread of activation of associations to other knowledge in the reader's memory to make these inferences (McNamara & Magliano, 2009). To put it differently, the information that is coming in reminds the reader of other knowledge they have. Processes that enable the reader to evaluate whether the new information can be fitted into their existing understanding of the world (e.g. Cook & O'Brien, 2014) follow this spread of activation. Thereafter the reader integrates the new information with their already existing knowledge to build a stronger knowledge representation of that specific topic, now including what they just read. Thereby, the knowledge representation becomes interconnected. This updated knowledge representation is then used when the reader proceeds to the next sentence, or chapter, to connect the new text with previous text and with the reader's previous knowledge (e.g. van den Broek, Young, Tzeng, & Linderholm, 1999). In this chain of reading processes, inference generation plays an important part in updating the understanding of the text and connect the text parts to create a whole. Without drawing on previous knowledge and generating inferences, readers end up with a superficial understanding of the text in which they understand the meaning of words, however, may not see how different parts of the text are connected. When generating inferences readers gain a deeper text understanding and build an interconnected representation of the text, i.e. a situation model (Johnson-Laird, 1983; van Dijk & Kintsch, 1983).

There are different types of inferences and while reading full texts, each inference the reader makes will affect the next. Together the inferences affect the offline understanding of the text (Hyönä, Lorch, & Kaakinen, 2002; van den Broek et al, 1999). A few of the most common inferences studied are text-connecting, knowledge-based, and predictive inferences. Text-connecting inferences refer to backward inferences by which the reader connects the focal text with the previous text. Knowledge-based inferences refer to inferences where the reader uses their own previous knowledge to fill in conceptual gaps of missing information. Predictive inferences refer to forward inferences by which the reader makes a prediction of the upcoming text. These inferences have often been studied by examining single inferences in isolation of other reading processes that are involved when the reader is making sense of a text, usually these studies also focus on reading a few sentences. However, various types of inferences are used in combination when the reader proceeds through a full text (Hyönä, Lorch, & Kaakinen, 2002; McNamara & Magliano, 2009). Because upper elementary school children learn content knowledge from reading texts, not only a few sentences, we studied children's inference generation while reading full texts. Similar studies concerning children's inference generation in full text have identified subgroups of readers that differ in the online processes they use (e.g. McMaster et al., 2012; Kraal et al., 2017). These studies focused on groups of developing readers with either a high or a low level of reading comprehension. In the current dissertation we aim to investigate subgroups of readers in a more heterogeneous group that is likely to have a larger resemblance to classroom populations. By studying differences in the number and type of inferences children produce we get a better understanding of how they build their situation model when reading and how these efforts relate to good offline reading comprehension.

1.2.2 Structuring Events When Building a Situation Model

To attain a coherent situation model of the text, the reader keeps track of and updates several dimensions of events unfolding in the text, such as when and where they play out, and what caused these events (e.g. van den Broek, 1990; Zwaan, et al., 1995). The ability to flexibly update one's understanding of the temporal order of two or several events in the text is important to continuously build an accurate situation model as the text unfolds. For example, a sentence like "Before you add or subtract a number, you should solve the multiplication" instructs the reader how to calculate correctly and is crucial for a child to understand in their math education. Research on understanding temporal relations of events has provided information about readers' ability to manipulate and update their situation model. Therefore, a great deal of research has been devoted to children's (e.g. Blything, Davies, & Cain, 2015; Cain & Nash, 2011; Natsopoulos & Abadzi, 1986; Pyykkönen & Järvikivi, 2012; Van Silfhout, Evers-Vermeul, & Sanders, 2015) and adults' (e.g. Münte, Schiltz, & Kutas, 1998; Ye, Kutas et al, 2012) understanding of temporal connectives such as *before* and *after*. The research field has seen different, sometimes competing, conclusions depending on the age of participants and the measurements used. Conclusions have differed on whether working-memory capacity or working-memory updating is taxed by these sentences. Conclusions have also differed on what textual demands are difficult for readers' to understand; grammatical complexity or the chronology of the text. Children's ability to understand temporal relations has often been researched in younger children using offline comprehension listening tasks or older children using reading tasks, whereas studies in adults have often used online reading tasks but with a limited set of materials. We aim to provide more clarity to differing conclusions by using extended measurements. We study reading processes of temporal relations in 9-12-year olds because it allows us to achieve important insights into both online and offline comprehension processes, in a period of life when language comprehension development interacts with the development of working memory. Because the ability to correctly comprehend temporal relations is important, especially in educational contexts, the seemingly contradictory findings in previous studies on this topic should be resolved, and the role of working memory should be clarified.

1.2.3 Building a Situation Model in Working Memory

Models of working memory generally agree that this is a system for the temporary storage and manipulation of information (for an anthology on models of working memory, see Miyake & Shah, 1999). Therefore, working memory is essential for situation model building (Daneman & Carpenter, 1980; Daneman & Merikle, 1996). Working memory is one of what are called executive functions in cognitive science (e.g. Diamond, 2013). Executive functions is an umbrella term for several general mental processes that help coordinate actions and control behavior to stay on task and go about daily life. An everyday example is to read and execute the instructions of a recipe when cooking. Although executive functions allow control of behavior, they come with limitations. For example, the storage capacity of working memory is limited. Consider all the details and bits of information that are contained in a single book; it is impossible for a reader to, at any given moment, keep every detail in a book active in working memory. Instead, concepts in the text are fluctuating in how active they are in the reader's mind while reading (e.g. van den Broek et al., 1999). The limitations of working memory can fluctuate within a reader but have also proven to differ between individuals and to be related to their reading comprehension. For example, children with a larger working memory capacity have better text comprehension (e.g. Cain, Oakhill, & Bryant, 2004; Christopher et al., 2012; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000; Swanson, Zheng, & Jerman, 2009).

Much of research on the role of working memory in developing readers has focused on the limitations of working memory capacity. However, theoretical models based on adult readers' reading comprehension not only entail a limited-capacity space that holds information in a heightened state of availability, but also considers the processes necessary to update the contents of this workspace. These processes aid readers as they continuously modify their mental representations of the text (Linderholm, Virtue, Tzeng, & van den Broek, 2004; McNamara & Magliano, 2009; van Dijk & Kintsch, 1983). Through a process of spread of activation (e.g. Collins & Loftus, 1975; Linderholm et al., 2004), concepts from the text and/or long-term memory that are automatically activated above a certain threshold may enter the focus of the reader's attention, i.e., working memory (van den Broek et al, 1999). Information useful to the reader is integrated in the situation model whereas information deemed not important becomes inhibited or forgotten (e.g. Kintsch, 1994; Wylie et al., 2018). Spread of activation is an efficient way of building comprehension because it consumes little of the reader's cognitive resources. However, syntactically or semantically complex sentences, requiring integration and sequencing of multiple pieces of information, often demand more than automatic processes. These types of sentences often involve strategic processing, i.e. reader-initiated processes, by which the reader consciously evaluates what is important (e.g. van den Broek & Helder, 2017). Keeping several units of information active in mind, while simultaneously trying to manipulate and analyze these pieces of information may be overtaxing children's developing cognitive resources (e.g. Blything et al., 2015).

The exact role of various working memory processes during text comprehension in development is still unclear (e.g. Kidd, 2013). This is partly due to controversies regarding the definition and operationalization of working memory (Cowan, 2017) as well as the attention that is given to storage components and to processing components in working memory. For typically developing readers, research has predominantly focused on storage and processing capacity similar to the span task addressed in the seminal paper by Daneman and Carpenter in 1980 (e.g. Seigneuric et al., 2000). A few examples also measure other executive functions such as inhibition (e.g. Christopher et al., 2012). However, to my knowledge, reading research generally does not include working memory tasks in which participants have to update their mental model of new information in a way that would simulate working memory processes as defined by models on reading comprehension (also noticed by Carretti, Cornoldi, De Beni, & Romanò, 2005).

The exact role of working memory in reading comprehension is likely to depend on the type of reading task. In a meta-analysis, Daneman and Merikle (1996) showed that working memory better predicted reading tasks that were specific (i.e. reading shorter pieces of text to understand the referent of a pronoun, to make a certain inference, or to revise inconsistencies) compared to more global text comprehension (i.e. reading a full text and drawing conclusions, and answering questions about the author's intent). Global comprehension of a full text indeed entails a wealth of different processes taking place in working memory. Therefore, if wanting to understand cognitive processes that underlie a certain reading task, researchers need to consider what aspect of working memory is believed to predict performance in specific reading processes. Researchers need to have specific hypotheses of why a certain working memory measurement is related to a specific reading task to optimally explain the relationship between the two. For example, a reading task when the reader needs to hold information in working memory for a certain period until resolving a reference, and a reading task when the reader needs to continuously update a cycle of events require comparison to different working memory tasks. To summarize, a limited capacity and processes for updating and modifying the content of the mental representation, influence text comprehension (Daneman & Carpenter, 1980; Daneman & Merikle, 1996). However, there is a need to clarify the relation between reading and these two different aspects of working memory in developing readers.

1.3 The Text

Texts vary on multiple levels, on macro levels such as discourse genre and text structure, and on micro levels such as sentence difficulty including word frequency and syntax complexity. Children's reading comprehension is predicted by their text structure knowledge, and interventions on text structure knowledge improve comprehension (e.g. Bogaerds-Hazenberg, Evers-Vermeul, & van den Bergh, 2020; Meyer, 1987; Meyer & Ray, 2011). Adults' guidance in choice of reading material can enhance the possibility of a positive challenge in children's reading advances and may sometimes be crucial to understanding the content (e.g. Snow & Rand, 2002). Knowledge on how children build their situation models in different texts helps educators when they choose, or help children choose, reading material. Knowledge on how children build their situation models in different texts is also useful when helping them find strategies suitable for the specific text types (Meyer, 1987). In the current dissertation, we study children's ability to build a situation model in texts that differ on macro levels and on micro levels.

1.3.1 Macro-Level Variation: Text Genre Differences

Children are generally exposed to narrative texts when learning to read and encounter expository texts later in their elementary-school years. Longstanding research of differences in text structure in narrative and expository texts shows that elementary school children find expository texts more difficult and explicit instruction is very much needed (e.g. Williams & Pao, 2011). Hence, when educators choose expository texts with the goal of teaching new content, they need to pay attention to the reading processes children already utilize and actively instruct on reading strategies to enhance learning (e.g. Lorch, 2015; Williams & Pao, 2011). For example, because expository text often is educative and explains new concepts, it is less likely that readers are able to connect the focal sentence with previous text unless explicitly instructed to generate these kind of backward inferences (Noordman, Kempf, & Vonk, 1992). In secondary-school students, inference generation is affected by reader proficiency, text accessibility and text genre (Denton et al., 2015). In college students, inference generation occurs less often for expository text than for narrative text (Lorch, 2017). To summarize, converging evidence shows that reading processes are, in part, modulated by text genre across childhood and early adulthood. By including various text genres in reading research, we are able to understand to what extent children's approach to text is stable and dependent on reader specific factors and to what extent their approach to text is dependent on text input.

1.3.2 Micro-level variation: Sentence level differences

On sentence and word levels, the reader uses text cues to build and update a situation model. Examples of such cues are temporal connectives that determine the order of events and causal cues that help the reader see the relation between events. Consider the sentences "*Before* you add or subtract a number, you should solve the multiplication", and "You should add or subtract a number *after* you solve the multiplication". When reading these two-clause sentences representing two events in a certain temporal order, readers use the temporal connectives *before* or *after* to understand in which order the two events occurred. However, micro-level variations in these cues may also impose processing costs that affect comprehension negatively. Research in pre-school children (Blything & Cain, 2016; Clark, 1971) and adults (e.g. Münte, Schiltz, & Kutas, 1998) shows that temporal connectives are not equally helpful. The connectives can impose certain demands that the reader needs to process, which can lead to a worsened reading comprehension. For example, conflicting ideas have been proposed by previous research regarding the position of the temporal connective. Sentence-initial-connectives may place a higher load on working memory capacity because the reader needs to hold that information active while reading the rest of the sentence (Blything & Cain, 2016). However, sentence-medial connectives may impose a working memory processing cost, because the reader needs to update the situation model they are building midsentence (Pyykkönen & Järvikivi, 2012). We set out to disentangle demands imposed by these micro level variations that can impede 9-12-year-old readers' comprehension of sentences containing temporal connectives.

1.4 The Reading Task

Reading is usually done with a certain purpose, be it for enjoyment, studying, or directly applying the information, such as using a cooking recipe. Because of variations in purposes and tasks connected to the reading experience, comprehension of the same text can be different. In other words, the task driving the reading scenario influences what the reader comprehends (Snow & RAND, 2002). For example, readers that do not routinely generate inferences may do so when prompted (Noordman et al., 1992). Task instructions, task demands, and how texts are presented are important aspects to consider for both scientific and educational practice. Methodological choices in a scientific study mean that we steer our understanding of the studied phenomena in a certain direction (van den Broek et al., 1993). Similarly, choices of instructional material or assessment material influence children's abilities to learn from text (Kendeou, van den Broek, Helder, & Karlsson, 2014). In both contexts, comparing various methods and materials yield a more complete picture of reading comprehension. In the current dissertation, we study readers' ability to comprehend text in various tasks.

1.4.1 Online and Offline Tasks

Cognitive reading models suggest that the online processes that the reader utilizes during reading will largely determine the outcome of the offline understanding of the text. For example, engaging in inference generation during reading causes deep comprehension; after an inference training, children improve their reading comprehension (Elleman, 2017). However, not in all instances do online processes seem to result in improved or changed offline comprehension (e.g. Rapp & Mensink, 2011). Online task instructions and online task demands may cause the reader to focus on a certain aspect of the text in a moment-to-moment text representation, whereas offline tasks may offer the possibility to focus on other or several aspects of the text in hindsight. In addition, readers may be unable or even unwilling to shift focus or reinterpret text during reading, whereas after having read the whole text they may do so. Including online and offline tasks clarifies how readers build a situation model of the text and may inform the research community of how methodological choices may interact with reader characteristics (e.g. Rapp & Mensink, 2011). Similarly, educational systems rely on instructing and assessing students using both online and offline reading tasks. Therefore, it is important to include both online and offline measures to fully apprehend how developing readers' text comprehension is formed.

1.4.2 Reading and Listening Comprehension

Although we aim to measure a certain type of comprehension process, different measurements may impose different demands on the reader. Thereby, different conclusions may be drawn that are not only related to the reading process of interest, but also related to individual differences in working memory (van den Broek et al., 1993). For example, contradicting hypotheses state, on the one hand, that sentence-initial connectives are more demanding for working memory and, on the other hand, that sentence-medial connectives are more demanding for working memory (Blything & Cain, 2016; Pyykkönen & Järvikivi, 2012). As these studies discuss different aspects of working memory, namely working memory capacity and working memory updating, they call for studying these sentences in relation to both these working memory measures. However, the predictions from the contradicting research stemmed from comprehension measurements that differed in which modality they were recorded, listening and reading, each potentially posing different demands on working memory. For example, decoding letters and words in a text may tax readers' working memory and in some ways these processes distract from understanding the meaning of the message, especially in developing readers (Kendeou et al., 2014). However, readers can control the speed with which they are encoding the written message. This allows them to slow down at difficult words or reread the text if necessary. These strategies can be vital in freeing up space in working memory for further comprehension processes. A listening task may be considered less demanding because listeners can concentrate on the message itself rather than on decoding letters and words (Kendeou et al., 2014). However, listeners do not have the opportunity to control the speed with which the message is conveyed, this depends on the speaker. Hence, listeners' working memory is taxed because they must try to keep the majority of the message active in working memory to comprehend the whole content. To better understand contradicting hypotheses and results from research using different presentation modalities, there is a need to map similarities and differences in comprehension of written and spoken communication.

In light of so-called Edtech advances, research on differences and similarities in comprehending written and spoken text is needed also to inform educational practices. In schools using digital learning platforms, 20% of the students are estimated to listen to digital read-aloud options as a help to understand text content (Magnusson Amu, 2020). The read-aloud options are popular amongst students; however, Edtech developers and researchers request more research of possible benefits of these practices (Grunér, Östberg, & Hedenius, 2018; Magnusson Amu, 2020; Wood, Moxley, Tighe, & Wagner, 2018). Because of these educational developments and the fact that working memory is still developing in older elementary school children, examining the role of working memory in comprehension in reading and listening tasks is valuable.

1.5 Outline of this Dissertation

The overarching aim of this dissertation is to examine cognitive reading comprehension processes in upper elementary school children. In doing so we study different situation-model building processes and how they are related to reader, text, and task characteristics. In the **second chapter** we study how children (9-11 years old) differ in online inference generation, and how these differences relate to text genres, and children's underlying reader characteristics. Therefore, we use a think-aloud task that allows readers to freely comment on the text while reading. From these data, we identify reader profiles that differ in both number and types of online inferences generated. This is done in both narrative texts and expository texts to understand whether reader profiles are stable traits or if children's inference generation changes depending on text genre. Furthermore, to understand underlying reader characteristics that may explain possible reader profiles, we examine the children's word-decoding ability, general text comprehension ability, non-verbal reasoning ability, working memory, and vocabulary. In the **third chapter** we study how differences in online inference generation relate to children's (9-11 years old) offline text memory. Therefore, we investigate whether children in the reader profiles identified in the second chapter, differ in their ability to structure their offline memory representation. To do so, we examine whether children show a centrality effect in their recall of the texts used for the think-aloud task. A centrality effect means that the reader remembers more of the central information, the gist, than of the details of the text. We also examine whether relations between the online and offline performance are qualified by text genre, i.e. narrative and expository texts. In the **fourth chapter** we examine children's (9-12 year old) ability to use temporal connectives when building a situation model of the text. In doing so, we examine the influence of text features such as the sentence position and familiarity of connectives, and clause salience. By means of two experiments, we aim to disentangle the effects that connective familiarity and clause salience have on comprehension. Importantly, we investigate how comprehension of sentences that include temporal connectives is qualified by children's working memory and use both a working memory capacity task and a working memory updating task for this aim. Testing both working memory capacity and working memory updating allows us to examine contradicting findings of earlier studies. Previous studies with contradicting findings did not only differ in their conclusions, but also in methodological choices to gather data, with a listening or with a reading task. In the **fifth chapter** we examine whether two different modalities, reading and listening, are taxing the comprehenders' working memory differently. As a first step to do so, we study adult readers' ability to make predictive inferences while reading and listening and included a working memory task. Finally, in the **sixth chapter**, results and conclusions from the four empirical studies are discussed in a broader context related to the important elements: reader, text, and task characteristics.

Chapter 2

Profiles of Young Readers:

Evidence from Thinking Aloud while Reading Narrative and Expository Texts

This chapter has been published as:
Karlsson, J., van den Broek, P., Helder, A., Hickendorff, M., Koornneef, A., & van Leijenhorst, L. (2018). Profiles of young readers: Evidence from thinking aloud while reading narrative and expository texts. *Learning and Individual Differences, 67,* 105–116. doi:10.1016/j.lindif.2018.08.001

Profiles of Young Readers:

Evidence from Thinking Aloud while Reading Narrative and Expository Texts

CHAPTER 2

Abstract

This study aimed to identify reading behavior profiles in nine-to-eleven-year-old children based on their think-aloud responses while reading narrative and expository texts. Three profiles emerged while reading narratives: Literal Readers, who stay close to the literal text by predominantly repeating it; Paraphrasing Readers, who extract meaning from the text by paraphrasing it; and Elaborating Readers, who use background knowledge to explain the text by generating inferences. The three profiles also emerged while reading expository text. Children generally exhibited the same profiles across the two text genres, however, expository texts elicited fewer correct inferences but more invalid inferences than did narratives, suggesting that children are influenced by text demands. Elaborating Readers had better word decoding skills, reading comprehension ability, and non-verbal reasoning ability than readers of the two other profiles, indicating a positive relation between inference generation and language abilities and cognitive resources.

Key words: Reading Profiles, Text Genres, Think Aloud, Developing Readers, Latent Profile Analysis

2.1 Introduction

Readers engage in various reading processes to understand a text. Importantly, readers need to go beyond the literal text and draw upon background knowledge to make inferences to understand the meaning of the text (e.g., Kintsch, 1988; van den Broek, 1990). Young readers differ in their ability to go beyond the text and generate necessary inferences (Cain & Oakhill, 1999; Kendeou, van den Broek, Helder, & Karlsson, 2014; Nation & Snowling, 1997). Such differences may result in children approaching texts in different ways. Studies using think-aloud procedures during reading of narratives revealed contrasting profiles in poor or good comprehending readers; readers in one profile stay close to the literal text, and readers in the other profile generate elaborative inferences that go beyond the text (Carlson, Seipel, & McMaster, 2014; Kraal, Koornneef, Saab, van den Broek, 2017; McMaster et al., 2012; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007; Seipel, Carlson, & Clinton, 2017). The identification of such reading profiles has led to the development of targeted reading interventions for poor comprehenders (McMaster et al., 2012). However, because reading comprehension is a multidimensional ability, it is important to consider reading profiles in developing readers across the whole range of reading comprehension ability, not only in poor *or* good comprehenders. Furthermore, because narrative and expository texts differ in text demands, it is important to compare readers' profiles for narrative texts to their profiles for expository texts. Considering reading profiles in different text genres may provide useful perspectives for theoretical questions. For example, whether children have a certain stable set of reading abilities with which they process text in a similar way across different situations, or whether they are influenced by conditions such as different text demands. Furthermore, such expansions allow important insights for evidence-based reading instructions across a larger group of developing readers and across different text genres.

2.1.1 Comprehension Processes

A reader can attain different levels of comprehension for a text, ranging from basic to deep understanding. A well-known distinction between such comprehension levels has been proposed in the construction-integration model (van Dijk & Kintsch, 1983; Kintsch, 1988, 1994). In this model three different levels are discussed: the *surface* level, where the reader encodes literal words and phrases, the *textbase,* where the reader understands referential relations within the text, and the *situation model,* where the reader enriches the mental representation of the text by elaborating on it and integrating background knowledge. Although various models on inference generation have been proposed, a consensus has emerged that inferences are important for building a situation model of the text (for a recent overview see, O'Brien, Cook, & Lorch, 2015). A reader that uses appropriate and global level inferences is more likely to reach beyond the surface level understanding and gains a textbase and situation model understanding of the text (e.g., Goldman, McCarthy, & Burkett, 2015; Graesser, Singer, & Trabasso, 1994; Kintsch, 1994). In developing readers, the ability to make adequate inferences during reading is causally connected to good reading comprehension (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Lynch et al., 2008; McGee & Johnson, 2003). Relevant for the current paper, there are three broad types of inferences that contribute to an extended understanding of texts. First, *text-connecting inferences* enable readers to connect a focal event with an event previously mentioned in the text. Text-connecting inferences tend to be routinized in good readers (e.g., McKoon & Ratcliff, 1992; Olson, 1985), and are facilitated by large vocabulary and large working memory (WM) capacity (Singer, Andrusiak, Reisdorf, & Black, 1992). Second, *elaborative inferences* enable readers to connect the text with relevant background knowledge. Elaborative inferences allow for causal connections and are important to create a rich and coherent mental representation of the text (e.g., Graesser et al., 1994; Lynch et al., 2008). Sufficient word reading abilities and world knowledge are some reader characteristics that facilitate the production of valid elaborative inferences (McNamara & Kintsch, 1996; Rapp et al., 2007). Although poor comprehenders also may generate elaborative inferences, these inferences are more often invalid than those of good comprehenders (McMaster et al., 2012). Third, *predictive inferences* are produced when readers predict upcoming events. Predictive inferences are not as routine or critical as the two previously mentioned inferences but rather depend on the text being constraining enough (Cook, Limber, & O'Brien, 2001; Kaakinen & Hyönä, 2005; Klin, Guzmán, & Levine, 1999; van den Broek, 1990). Furthermore, the likelihood of making predictive inferences depends on the interaction of reader characteristics, such as WM capacity, and text characteristics, such as high causality between text parts (e.g., Linderholm, 2002). These three types of inferences may contribute differentially to young readers' ability to process text beyond the literal level and build an enriched mental representation.

Given the positive effects that inference processes have on reading comprehension, it is important to identify whether some children consistently process the text on a basic level whereas others are better able to enrich their mental representation using elaborative inferences. Although much research indicates that good readers generate more inferences than poor readers (e.g., McNamara & Kintsch, 1996), differences have also been found within poor comprehending readers (Rapp et al., 2007). Using a think-aloud procedure, two subgroups of nine-to-ten-year-old poor comprehenders have emerged (Carlson et al., 2014; McMaster et al., 2012; Seipel et al., 2017). One subgroup of children stayed close to the basic meaning of the text, mainly repeating or paraphrasing the text (Paraphrasers). The other subgroup of children used background knowledge to make inferences, albeit sometimes erroneously (Elaborators). Similar reading profiles have been found in a younger group of Dutch poor comprehending readers, and in their good comprehending peers (Kraal et al., 2017), indicating promising generalizability of reading profiles. Continuing research of reading profiles may help to better understand whether children have a certain approach to process text and whether that approach relates to different levels of text comprehension as described by influential reading models (e.g., Johnson-Laird, 1983; van Dijk & Kintsch, 1983). In the current study we make no a priori distinction between good and poor comprehenders, but aim to identify homogenous subgroups, characterized by their reading behavior, within a heterogeneous population spanning from poor to good comprehension abilities.

2.1.2 Text Genres

Expository texts are often more difficult than narratives for developing readers (e.g., Best, Floyd, & McNamara, 2008), and several reasons may explain differences in text demands (e.g. Eason, Goldberg, Young, Geist, & Cutting, 2012). First, topics and hence familiarity of words may differ between the two text genres. Narratives often include everyday language whereas expository texts often introduce new words and terminology (Medina & Pilonieta, 2006). Therefore, expository texts often have a higher information density. Second, compared to narratives, expository texts are often more varied with regard to their structure (e.g. Lorch, 2015). Narratives often follow a more or less similar structure with similar elements and timelines (such as the protagonists initiating goal, actions, reactions, and outcomes; e.g. Mandler & Johnson, 1977). Expository texts come in different formats and require the reader to apply more varied reading strategies (Lorch, 2015). For example, there is not necessarily a timeline to follow but readers need to understand several subordinate ideas in relation to a main idea (Meyer, 1987). Hence, it is important to understand whether children approach the two text genres differently. Young readers are likely to lack in knowledge of both topic (e.g., Samuelstuen & Bråten, 2005) and text structure (e.g., Williams, Hall, & Lauer, 2004), making it difficult to effortlessly comprehend expository texts. For these reasons, expository texts pose difficulties in making inferences using background knowledge, especially for readers who already lag behind in comprehension skills. Indeed, children with poor inferencing skills experience comprehension difficulties when reading expository text (e.g., Best et al., 2008; Eason et al., 2012; Kraal et al., 2017; Schellings, Aarnoutse, & van Leeuwe, 2006). In adolescents, poor readers generate fewer inferences while reading expository compared to narrative texts (Denton et al., 2015). However, strategically elaborating on expository texts, if anything, *facilitates* in-depth comprehension of expository texts and, therefore, an *increase* in inference making would be desirable (Lorch, 2015; Mayer, 1996). By comparing inference skills in reading profiles of elementary school children across narrative- and expository texts we may examine whether developing readers recognize different text demands, and identify whether children with a certain reading profile could benefit from more practice with inference generation while reading expository texts.

2.1.3 Reader Characteristics

Because text comprehension is a multidimensional ability, different reading profiles may be related to individual differences in other language abilities and cognitive resources. Individual differences predict reading comprehension in both adult and developing readers (e.g., Hannon, 2012; Language and Reading Research Consortium, & Logan, 2016). In particular, and as mentioned above, good word decoding, reading comprehension skills (e.g. Carlson et al., 2014; Olson, 1985; Rapp et al., 2007), large WM capacity, and vocabulary promotes the ability to make different types of inferences while reading (Linderholm, 2002; Linderholm & van den Broek, 2002; Singer et al., 1992). However, some children's inference problems may be caused by a limited vocabulary (Nation & Snowling, 1998, 1999), whereas others struggling with inference generation may possess enough lexical knowledge but not know how to draw on this knowledge (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999), possibly due to an immature reasoning ability (de Leeuw, Segers, & Verhoeven, 2016; Naglieri, 2001). When tracing the heterogeneity in developing readers back to a number of underlying homogeneous reading profiles, it is important to also map out whether these profiles differ in word decoding, reading comprehension skills, vocabulary, non-verbal reasoning skills, and WM capacity to better understand underlying competences.

2.1.4 Distinguishing within Processes Close to Literal Text

So far, we have focused on the importance of inference making beyond the textual information. However, while participating in think-aloud studies readers also show understanding for the literal meaning of text by producing a substantial number of *text repetitions* and *paraphrases*. Prior research on reading profiles has aggregated text repetitions and paraphrases (Carlson et al., 2014; Kraal et al., 2017; McMaster et al., 2012), but there are reasons to examine whether these contribute differently to young readers' literal and in-depth comprehension of text. Literal repetition of text is a superficial reading strategy used when more in-depth reading strategies are too demanding: readers with a low WM capacity more often resort to text repetitions at the expense of demanding inferences than readers with a high WM capacity (Linderholm & van den Broek, 2002). Paraphrasing a text has been described as a superficial reading strategy for the same reason; instead of making inferences readers paraphrase the text (Magliano & Millis, 2003; Seipel et al., 2017). However, paraphrasing a text has also proven useful to start up higher order comprehension processes, i.e. facilitating inference generation (e.g., Denton et al., 2015; McNamara, 2004). Rephrasing the concepts in a text into one's own words strengthens memory traces for the text (e.g., Bohn-Gettler & Kendeou, 2014) and makes the focal text and its semantic relations more comprehensible to the reader (Coté, Goldman, & Saul, 1998). Hence, distinguishing between text repetitions and paraphrases may contribute to our knowledge about how young readers differ in their ability to start up inferences to enable building a rich representation of the text.

2.1.5 Current Study

In the current study we use a think-aloud procedure to examine how nine-to-eleven-year-old children approach text. We aim to identify subgroups of children characterized by their profile of think-aloud responses while reading narrative and expository texts. Contrary to previous research focusing on poor or good comprehenders, we include readers across a larger range from poor to good reading comprehension ability. To trace back the heterogeneity in these readers to a set of underlying homogeneous subgroups we use a Latent Profile Analysis (LPA). Individual differences in a set of variables, such as categories of think-aloud responses, are complex and often both quantitative and qualitative in nature. Linear models are not always suitable to capture the complexity of such data. LPA is a powerful and flexible tool, able to identify individuals with similar response patterns in such complex sets of variables (e.g., Bergman & Magnusson, 1997; Hickendorff, Edelsbrunner, McMullen, Schneider, & Trezise, 2017).

First, in the current study, we compare reading profiles across text genres. Previous research shows that young readers have more reading difficulties with expository texts compared to narrative texts. Poor comprehenders, in particular, are less able to make inferences while reading expository texts. In the current study, we compare children's think-aloud responses to narrative texts with those to expository texts, by conducting separate LPAs and investigating the similarities in the profiles of reading behavior at the group level, and the interrelation between the profiles of individual children. Identifying reading profiles across different text genres helps to examine whether developing readers recognize and adjust to different text demands. Second, in light of previous research we anticipate that children differ in their ability to make inferences during a think-aloud task. On the one hand, research investigating reading profiles shows that this difference occurs within poor comprehenders. On the other hand, other studies on inference making show that good comprehenders generally are better able to generate inferences while reading than poor comprehenders. In the current study we are able to put these findings in relation to one another; by conducting MANOVA's we investigate whether the children with different reading profiles based on think-aloud responses differ in their language abilities (word decoding, reading comprehension, and vocabulary) and in their cognitive resources (non-verbal reasoning ability, and working memory). Third, methodologically we expand on previous research by distinguishing not only between far-from-text processes (i.e. inferences) and close-to-text processes, but also within processes close to the literal text: verbatim repeating and paraphrasing the text. The first may entail text decoding on a surface level whereas the latter may be important to start up inference-generation processes. Hence, we may examine differences in inference generation and also whether readers of different profiles differ in their close-to-text processes.

2.2 Methods

2.2.1 Participants

One hundred and seven children (61 girls) between nine and eleven years old (*M* = 10.3, *SD* = 0.73) participated in the current study. Children from grades six and seven were recruited from 16 different schools in the south-west of the Netherlands. Parents gave written consent for participation and children gave oral consent. Participants' consent was obtained according to the Declaration of Helsinki, and the study was approved by the Ethical Committee at Leiden University. Inclusion criteria were having a typical development and Dutch as native language. In addition, only children scoring above the 40^{th} percentile on the Dutch national standardized word reading test, the *Three Minute Test* (CITO Drie Minuten Toets, (DMT), Krom, Jongen, Verhelst, Kamphuis & Kleintjes, 2010), were included to ensure they had average to good technical reading skills. To ensure that children with a range of comprehension proficiency were included from both grades, children of all proficiency levels on the Dutch national standardized CITO reading comprehension test (Feenstra, Kleintjes, Kamphuis & Krom, 2010; Weekers, Groenen, Kleintjes & Feenstra, 2011) were included.

2.2.2 Materials

2.2.2.1 Think-aloud Protocol

To assess text comprehension processes during reading we used a think-aloud protocol (e.g., Clinton & Van den Broek, 2012; McMaster et al., 2012; Van den Broek, Lorch, Linderholm, & Gustafson, 2001). Participants read two narrative and two expository texts. We used P-CLIB version 3.0 (Evers, 2008) to calculate basic text characteristics, and T-scan (Pander Maat, Kraf, & Dekker, 2017) to calculate text cohesion, see Table 2.1. Different measures on text length from P-CLIB and the Development Level (D-level) from T-Scan together indicate that syntactical complexity is fairly similar across text genre. (The D-level is a combined measure on syntax complexity which assigns sentences to a level on an 8-level scale, based on how difficult the sentence is from a developmental perspective; Pander Maat, Kraf, & Dekker, 2017). The percentage of frequent words and type-token ratio together indicate a higher information density of expository texts compared to narrative text, whereas the higher number of connectives per clause indicate higher text cohesion of expository text. The two narrative texts followed a typical story structure: in the beginning the protagonist had a task to complete (making or buying a gift for a family member) and made a few attempts before fulfilling the goal (e.g., Mandler & Johnson, 1977). The two expository texts followed a descriptive text structure: a nature phenomenon (earthquakes and snakes) was described with main and subordinate ideas (e.g., Meyer, 1987). English translations of the four texts are included in the supplementary file Translated Test Texts.

Table 2.1

Text Characteristics of the Narrative (N) and Expository (E) Texts based on P-CLIB[(a)] and T-Scan[(b)].

	Mieke picks a present (N)	Holiday shopping (N)	When the earth shakes (E)	Dutch snakes (E)
No. of sentences/ text[a]	17	16	15	16
Average no. of words/ sentence[a]	12	12.4	11.7	11.4
Average no. of letters/ word[a]	4.6	4.5	5.3	4.6
Percentage of frequent words[a]	75.61	77.39	72.57	74.18
Grade level[a]	7	7	8	7
D-level (proportion sentences > level 4)[b]	0.18	0.25	0.27	0.19
Type-Token ratio[b]	0.48	0.56	0.67	0.62
Connectives per clause[b]	0.13	0.11	0.28	0.35

Two undergraduate students (both female) and the first author collected the data, hereafter named test leaders. Before collecting the data, the test leaders set up a test protocol and practiced how to instruct the participants for each task. In the think-aloud task, each sentence was printed on a separate page. Participants were instructed to read each sentence aloud and report what they were thinking before moving on to the next sentence. Instructions were explicit in mentioning that there were no correct or incorrect responses, but that the test leader was just interested in hearing what came to their minds when reading the text. Before participants read the four test texts, they received a narrative practice text. For the first half of this text, the test leader modelled the think-aloud procedure following a script to ensure all participants received the same examples (including paraphrases, comments about the text, and different kinds of inferences), i.e. the test leader read a sentence and then modelled thinking aloud about the sentence. Then the participant practiced with the second half. During reading of the four test texts there was no feedback, only neutral encouragement to think aloud such as “What are you thinking now?” when the participant did not respond to the sentence he or she had read (e.g., Pressley & Afflerbach, 1995).

The recorded think-aloud session for each participant was transcribed and categorized by the undergraduate students while supervised by the authors. For each sentence, the think-aloud responses were parsed into subject-verb clauses (idea units) and assigned to categories based on previous research (e.g., Linderholm &

van den Broek, 2002; McMaster et al., 2012; Trabasso & Magliano, 1996). The inter-rater reliability of the two undergraduate students was good, $K = 0.75$, $p < .01$. The following six categories are relevant to the current study (see Table 2.2): text repetitions, paraphrases, text-connecting inferences, valid elaborative inferences, invalid elaborative inferences, and predictive inferences.

Table 2.2

Labels and Description of Think-aloud Response Categories

Response category	Description	Example She decided to make a necklace for her mom
Text Repetition	Literal repetition of the text	She decided to make a necklace for her mom
Paraphrase	Repeating the text in own words	She had made up her mind, she would put together a pretty necklace to give her mom
Text-connecting Inference	Connecting and reinstating events from prior text	A similar one that her friend made for her mom
Valid Elaborative Inference	Correct use of prior knowledge to explain the text	Because parents like when their children make them things
Invalid Elaborative Inference	Incorrect use of prior knowledge to explain the text	So she had to go to the store to get one
Predictive Inference	Predicting upcoming events in the text	I think her mom will be very happy when she opens the present

Only the first categorized response unit, i.e. the first idea unit the participant responded with, for each sentence was used in the analyses. The initial response is thought to indicate spontaneous thoughts in response to the text itself, whereas subsequent responses may reflect other, interfering thoughts (e.g., Ericsson & Simon, 1980; Hertzum, Hansen, & Andersen, 2009). Across the two narrative texts, the number of initial responses was averaged for each response category (e.g. for text repetitions, for paraphrases and so on) for each participant. Thereafter, a percentage score was calculated for each response category in relation to the total number of responses for this text genre. Similarly, initial response averages and percentages were calculated for each category and child across the two expository texts. Responses such as meta-cognitive or emotional comments, questions, invalid responses to other categories, and inaudible segments were coded but each made up a small percentage of all responses and were removed from further analyses.

2.2.2.2 Word Decoding

As a standardized measure of word decoding skills, we used the normed ability scores of the CITO DMT (The Three Minute Test developed by the Dutch Central Institute of Test Development; Krom et al., 2010) provided by the school. The reliability of this test is good, $\alpha > .921$ (Krom et al., 2010). The DMT is a word decoding test on which participants have to read aloud words without context as fast and as accurately as possible. The scores of the participants ranged from 80 to 124, reflecting average to good word decoding ability.

2.2.2.3 Reading Comprehension

As a standardized measure of reading comprehension skills, we used the normed ability scores of the Dutch national CITO reading comprehension test (Feenstra et al., 2010; Weekers et al., 2011) provided by the school. The reliability of this test is good, MAcc > .89 (Feenstra et al., 2010; Weekers et al., 2011). This is a paper-and-pencil test where the scores are based on how well the participant answers multiple choice and open-ended questions after having read narrative and expository texts. The questions tap into understanding factual information in the texts, inference ability - both within the text and with prior knowledge, and knowledge about text structure. The multiple-choice questions have four answering possibilities. The scores of the participants ranged from 11 to 98, reflecting the whole range of reading comprehension ability.

2.2.2.4 Vocabulary

To assess receptive vocabulary, we used a Dutch version of the Peabody Picture Vocabulary Test adapted for group administration (PPVT: Schlichting, 2005). The reliability of the Dutch PPVT is good, $\lambda\text{-}2 > .89$ (Schlichting, 2005). Participants received a booklet with one word on each page. Under each word four pictures were displayed. To show conceptual knowledge of the word, participants were asked to circle the picture that matched the word. Each individual's score was the number of correctly circled pictures within 15 minutes. There was one practice item and 60 test items, the highest possible score was 60. The scores of the participants ranged from 23 to 56 points.

2.2.2.5 Reasoning Ability

To assess non-verbal reasoning ability, we used a group-administered version of Raven Standard Progressive Matrices (Raven SPM: Raven, Raven, & Court, 1998). The reliability for Raven SPM is good, > .98 (Raven, 2008). Participants received a booklet with matrices with a missing part and were instructed to "solve as many puzzles" as possible within 30 minutes. The participants' task was to find the missing part out of six or eight options. For each item they wrote the number of the correct answer on an answering sheet. The items continuously increase in difficulty throughout the test. An individual's score was the number of correct answers. There was one practice item and 59 test items, the practice item was included in the score and the highest possible score was 60. The scores of the participants ranged from 13 to 53 points.

2.2.2.6 Working Memory

To assess verbal working memory (WM) we used a Dutch version of the Sentence Span task (Swanson, Cochran, & Ewers, 1989). In this task both manipulation of information (answering a question) and maintenance of information (remembering words) are measured. At the first level, the test leader read two unrelated sentences and thereafter asked an open-ended question about the content of one of the sentences. Participants were instructed to first answer the question and then say aloud the last word of each of the two sentences. For each subsequent level, WM load is increased by adding one sentence until reaching the final level with five sentences. There were three practice trials on the first level and two test trials for each level. If a participant made an error on both trials within a level, the test was discontinued. An individual's score was the number of correctly remembered words for trials where the question was answered correctly. This scoring method has gained a good internal consistency of .79 (Conway et al., 2005). The highest possible score is 28. The scores of the participants ranged from 1 to 22 points.

2.2.3 Procedure

Each child participated in two test sessions, an individual session (lasting approximately one hour) and a group session (lasting approximately 45 minutes). The tasks were administered in the same order to each participant during the individual session, first the think-aloud task and second the WM task. The think-aloud task started with the practice text. Thereafter the narrative and expository test texts were presented in an interleaved fashion; Mieke picks a present, When the earth shakes, Holiday shopping, and, finally, Dutch Snakes. There was a break before the WM task was administered. Group sessions were held with all participants within each school during one occasion. First the Raven SPM was administered followed by the Peabody Picture Vocabulary Test. After the individual session, the participants were thanked and received a small reward.

2.2.4 Analyses

To identify subgroups of children characterized by a profile of think-aloud responses while reading narrative texts, we performed a Latent Profile Analysis (LPA; Hagenaars & McCutcheon, 2002; Oberski, 2016), carried out in version 5.0 of the statistical program Latent Gold (Vermunt & Magidson, 2013). LPA is a model-based cluster analysis technique, aiming to identify qualitative individual differences (i.e., subgroups or clusters) based on individuals' scores on a set of continuous variables. The following five think-aloud categories were included as continuous variables: the percentage of text repetitions, paraphrases, text-connecting inferences, valid elaborative inferences, and predictive inferences. As there were only few observed values of the variable 'percentage of invalid elaborative inferences', this variable was included as an ordinal (also compatible with the LPA, Vermunt, Tran, & Magidson, 2008) rather than as a continuous variable. To select the optimal number of latent clusters we used statistical information criteria, Bayesian Information Criterion (BIC) and Consistent Akaike Information Criterion (CAIC), which represent a trade-off between model fit (log-likelihood) and model complexity (the number of estimated parameters), combined with the interpretability of the resulting clusters (e.g., Hickendorff et al., 2017). Entropy and classification measures were used to evaluate the absolute fit of the model to the data.

To compare the reading profiles across the text genres (narrative vs. expository text), we performed a second LPA on the think-aloud responses to the expository texts. For this analysis, the percentages of invalid elaborative inferences and of predictive inferences were included as ordinal variables, as these held few observed values. The remaining four think-aloud categories were included as continuous variables. To select the optimal number of latent clusters we used the BIC and CAIC, combined with interpretability of the resulting clusters. To test whether individual children have a similar response profile across narrative and expository texts, we performed a cross tabulation with a chi-square test for independence between the cluster membership based on narratives and the cluster membership based on exposition.

To examine whether the identified profiles are related to differences in language abilities and cognitive resources, we performed two one-way between-groups Multivariate Analysis of Variance (MANOVA). Children's performance scores on word decoding, reading comprehension, vocabulary, non-verbal reasoning ability, and working memory were entered as dependent variables and their LPA cluster membership (based on reading narrative or expository texts) as fixed factor.

Descriptive statistics and correlations of the variables used in the above-mentioned analyses are reported in the supplementary Tables 2.S1 and 2.S2, respectively.

2.3 Results

2.3.1 Profiles Characterized by Think-Aloud Responses to Narrative Text

To identify reading profiles based on narrative texts we conducted LPAs with one to eight clusters. The model with three clusters had the lowest BIC-value and CAIC-value. This model had an R-square entropy of .93 (values range between 0 and 1, and higher values indicate more certainty of classification; Collins & Lanza, 2010) and a classification error of .022, indicating that children's cluster membership could be predicted from their responses very well. The following five response types made significant contributions to the classification: text repetitions, paraphrases, text-connecting inferences, valid elaborative inferences, and predictive inferences (Wald's > 43, all *p*s <.001). Invalid elaborative inferences did not make a significant contribution to the classification (Wald-statistic = .14, $p = .93$). The profiles of the three clusters are visualized in Figure 2.1.

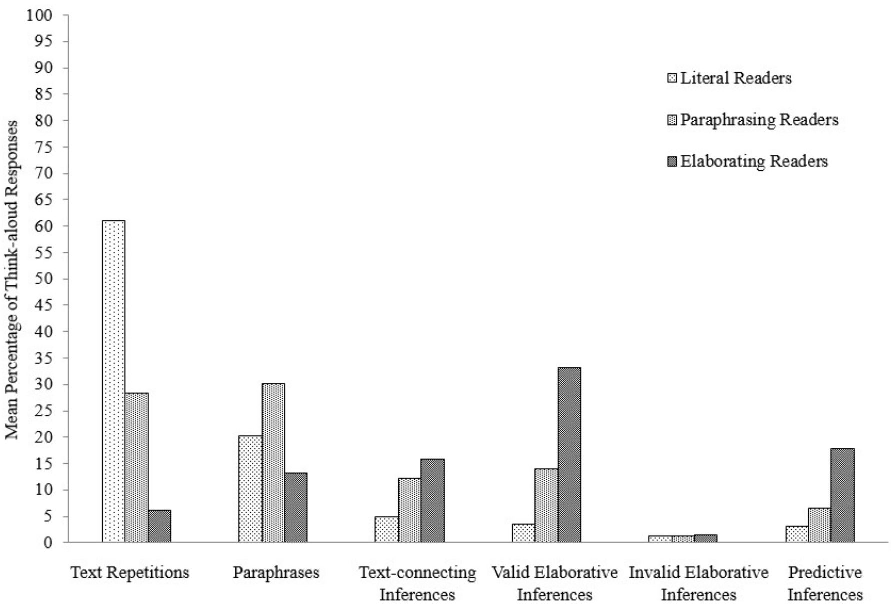

Figure 2.1. Clusters characterized by think-aloud responses while reading narrative texts. The clusters' mean percentage of responses is based on the LPA-estimates and shown on the y-axis. Type of think-aloud category is shown on the x-axis.

The three profiles seem to reflect different approaches to text processing based on the type of responses children in the three clusters were most likely to produce. The largest cluster (43.6% of the children) consisted of *Paraphrasing Readers*. Children belonging to this cluster primarily paraphrased the text, more so than the other two clusters (all *p*s < .008, see Supplementary Table 2.S3). In addition, they often repeated the text literally, and produced a fair number of text-connecting inferences and valid elaborative inferences. The second largest cluster (34.5% of the children) consisted of *Elaborating Readers*. Children belonging to this cluster primarily used valid elaborative inferences to explain the text. In addition, they produced a fair number of text-connecting inferences and predictive inferences. They differentiate from the other two clusters primarily in the number of valid elaborative inferences and predictive inferences (all *p*s < .001, see Supplementary Table 2.S3). The third cluster (21.9% of the children) consisted of *Literal Readers*. Children belonging to this cluster showed a large number of text repetition responses, more so than the other two clusters (all *p*s < .001, see Supplementary Table 2.S3). In addition, they produced a fair number of paraphrases.

2.3.1 Profiles Characterized by Think-Aloud Responses to Expository Text

To examine whether children approach narrative and expository texts similarly we first identified profiles of readers on their responses to expository texts by performing an LPA. From the LPAs with one to eight clusters, the three- and four-cluster models had virtually identical BIC values. Of these two, the three-cluster model had the lowest CAIC value, therefore we report the three-cluster model. This model had an R-square entropy of .87 and a classification error of .057. All response types made significant contributions in the cluster classification (Wald's > 6.8, all *p*s <.05). The profiles of the three clusters are visualized in Figure 2.2.

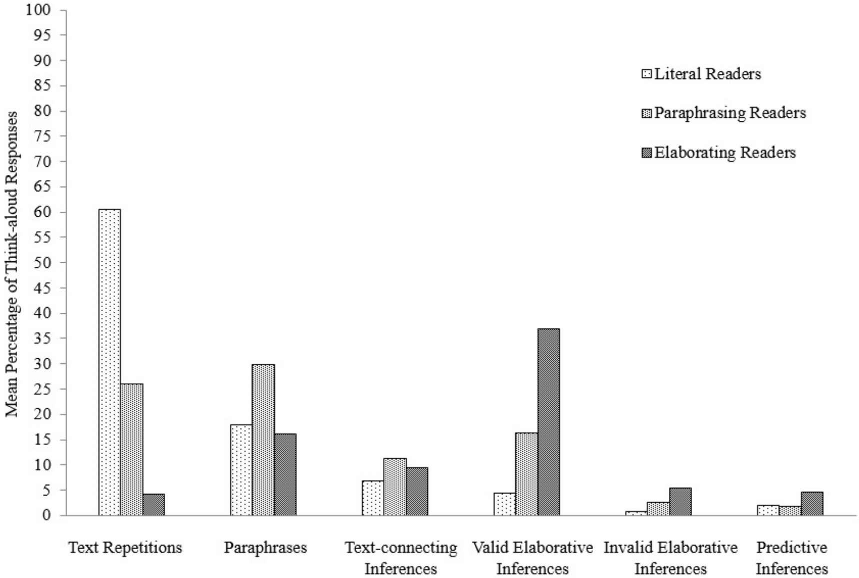

Figure 2.2. Clusters characterized by think-aloud responses while reading expository texts. The clusters' mean percentage of responses is based on the LPA-estimates and shown on the y-axis. Type of think-aloud category is shown on the x-axis.

The three profiles resemble the profiles identified in the LPA on narrative texts and reflect similar approaches to text based on their think-aloud responses. Paraphrasing Readers (47.5% of the children) primarily paraphrased the text or repeated it literally; in addition, they produced a fair number of text-connecting inferences and valid elaborative inferences. Literal Readers (26.9% of the children) produced a large number of text repetitions; in addition, they produced a fair number of paraphrases. Elaborating Readers (25.6 % of the children) primarily used valid elaborative inferences to explain the text; in addition, they produced a fair number of paraphrases and text-connecting inferences. Inspection of Figures 2.1 and 2.2 shows that the response patterns are very similar across narrative and expository texts for all profiles, but the number of inferences differ between the two text genres. Expository texts elicited fewer predictive inferences in both Paraphrasing Readers and Elaborating Readers than narrative texts. Furthermore, both Paraphrasing Readers and Elaborating Readers made more invalid elaborative inferences while reading expository than while reading narrative texts, with the largest difference within the Elaborating Readers. Elaborating Readers also produced fewer text-connecting inferences when reading the expository texts compared to narrative texts.

2.3.3 Comparison between Profiles Characterized by Narrative and Expository Texts

To examine whether individual children showed a similar reading profile across narratives and exposition we performed a cross tabulation of the classification based on the narrative and the expository text. First, all participants were assigned to a profile based on reading the narrative texts, thereafter, based on reading the expository texts. To do so a modal assignment procedure (Vermunt & Magidson, 2013) was used by which children were assigned to the profiles for which they had the highest posterior probability on narrative and expository text, respectively. Across all children, the average classification error was low, .022. The Kendall's tau-b showed a significant relation between profiles on narrative and expository texts, χ^2 (4, N = 107) = 80.43, $p < .001$, Kendall's tau-b = .71. The majority of Literal Readers (79.2 %), Paraphrasing Readers (71.7 %), and Elaborating Readers (64.9 %) had the same approach to both narrative and expository texts (see Figure 2.3). Across all profiles, 24.6% of the children had a different profile in which they used *fewer* inferences and more paraphrases and repetitions while reading exposition, compared to narratives. Only 8.5% of the children had a different profile in which they used *more* inferences while reading exposition, compared to narratives. The results suggest an ordinal relation between different profiles, which is supported by the high Kendall's tau-b. If an individual approached expository texts differently than narrative texts, there was indeed an ordinal manner in which such change in text approach occurred. For example, some individuals assigned to the profile Elaborating Readers while reading narratives became assigned to the profile Paraphrasing Readers while reading exposition, but none changed into Literal Readers. To summarize, children approach narrative and expository texts similarly and, if they use a different approach, they are likely to stay closer to the literal text in expository texts compared to narrative texts.

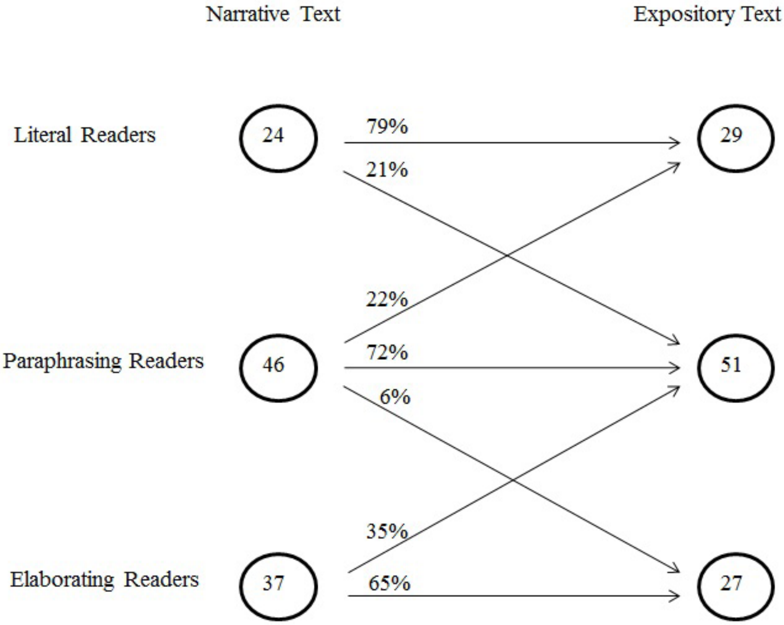

Figure 2.3. Children's profile membership for narrative and expository texts. In the circles, the total number of children in each cluster type is reported for each type of text. The arrows show the change in think-aloud response profiles; the percentages of children keeping the same profile or changing to a different profile are indicated above each arrow.

2.3.4 Profiles Related to Language Abilities and Cognitive Resources

Two MANOVAs were performed to examine whether the identified profiles are related to differences in word decoding, reading comprehension, vocabulary, reasoning ability, and working memory. Due to missing data in the dependent variables for three children the total number of children is 104 in these analyses. For the first MANOVA, all participants were assigned to their profile based on reading the narrative texts (a modal assignment procedure; Vermunt & Magidson, 2013). The data did not show any violations of assumption of homogeneity of variance-covariance matrices ($p > .001$). The equality of variance was not violated for any of the dependent variables (all *p*s > .05), except for word decoding, $p = .040$. As word decoding did not show equality of variance, we used the Pillai's Trace significance level for a more robust test. There was a significant difference between the three profiles on the combined dependent variables, F (10, 196) = 2.88, $p = .002$, $\eta_p^2 = .13$. Means and standard errors for each profile are reported in Table

2.3. Considering the dependent variables separately in five univariate ANOVAs with Bonferroni-corrected alpha levels of .01, cluster membership had a significant effect on word decoding, F (2, 101) = 6.79, p = .002, η_p^2 = .12, reading comprehension, F (2, 101) = 8.45, p < .001, η_p^2 = .14, and non-verbal reasoning ability, F (2, 101) = 7.36, p = .001, η_p^2 = .13. Neither vocabulary, F (2, 101) = 1.79, p = .171, η_p^2 = .03, nor working memory, F (2, 101) = 2.89, p = .060, η_p^2 = .05, differed significantly between the three profiles. Post-hoc analyses comparing mean scores between the profiles on the three dependent variables with a significant effect were conducted with Tukey HSD adjustment. The Elaborating Readers performed significantly better than Literal Readers on word decoding (p = .003), reading comprehension (p < .001), and reasoning abilities (p = .001). The Elaborating Readers also performed significantly better than the Paraphrasing Readers on word decoding (p = .012), reading comprehension (p = .039), and reasoning abilities (p = .019). Although Paraphrasing Readers numerically performed better than Literal Readers on all three measures, none of these differences were significant (all ps > .103).

Table 2.3

Means and Standard Errors from the MANOVA for Each Dependent Variable and Each Reading Profile Based on the Narrative Texts.

	Literal Readers (N = 23)	Paraphrasing Readers (N = 45)	Elaborating Readers (N = 36)
Word Decoding	93.35 (1.94)**	95.62 (1.40)*	101.69 (1.55)
Reading Comprehension	30.78 (3.59)***	39.89 (2.57)*	49.44 (2.87)
Reasoning Ability	34.70 (1.58)***	37.42 (1.13)*	42.08 (1.26)
Vocabulary	40.17 (1.36)	41.10 (0.97)	43.22 (1.09)
Working Memory	6.74 (0.96)	8.24 (0.68)	9.67 (0.77)

Note. Asterisks indicate where the Literal and Paraphrasing Readers differed from the Elaborating Readers at * p < .05, ** p < .01, and *** p < .001. There were no significant differences between the Literal and the Paraphrasing Readers.

For the second MANOVA, all participants were assigned to a profile based on reading the expository texts (Vermunt & Magidson, 2013). The data did not show any violations of assumption of homogeneity of variance-covariance matrices (p > .001) or of assumption of equality of variance for any of the dependent variables (all ps > .05). We kept using the Pillai's Trace significance level for a robust test. There was a significant difference between the three profiles on the combined dependent variables, F (10, 196) = 1.90, p = .047, η_p^2 = .09. Means and standard errors are reported in Table 2.4. Considering the dependent variables separately in five univariate ANOVAs with Bonferroni-corrected alpha levels of .01, cluster membership had a significant effect on non-verbal reasoning ability, F (2, 101) = 6.07, p = .003, η_p^2 = .11, and a marginally significant effect on reading comprehension, F (2, 101) = 4.76, p = .011, η_p^2 = .09. There were no effects on word decoding, F (2, 101) = 3.80, p = .026, η_p^2 = .07, vocabulary, F (2, 101) = 2.25, p = .111, η_p^2 = .04, or working memory, F (2, 101) = 1.29, p = .280, η_p^2 = .02. Post-hoc analyses showed that the Elaborating Readers performed significantly better than Literal Readers (p = .002), and Paraphrasing Readers performed marginally better than Literal Readers (p = .058) on reasoning abilities. There was no difference between Elaborating Readers and Paraphrasing Readers (p = .239) on reasoning abilities. In addition, the Elaborating Readers performed significantly better than Literal Readers (p = .008) on reading comprehension, no other differences were significant (all ps > .121).

Table 2.4

Means and Standard Errors from the MANOVA for Each Dependent Variable and Each Reading Profile Based on the Expository Texts.

	Literal Readers (N = 27)	Paraphrasing Readers (N = 50)	Elaborating Readers (N = 27)
Word Decoding	93.35 (1.94)	95.62 (1.40)	101.69 (1.55)
Reading Comprehension	34.22 (3.43)**	40.66 (2.52)	49.11 (3.43)
Reasoning Ability	34.52 (1.48)**	38.7642 (1.08)	41.74 (1.48)
Vocabulary	39.48 (1.25)	41.98 (0.92)	43.11 (1.25)
Working Memory	7.41 (0.90)	8.38 (0.66)	9.44 (0.90)

Note. ** indicate where the Literal Readers differed from the Elaborating Readers at p < .01. No other differences were significant.

2.4. Discussion

We examined possible reading profiles using children's think-aloud responses while reading narrative and expository texts and investigated whether young readers approach narrative texts and expository texts differently. Furthermore, we examined whether children with distinct reading profiles differ in language abilities and cognitive resources. In pursuing these aims, we made a distinction within close-to-the-literal-text processes, namely between text repetitions and paraphrases.

Three distinct reading profiles that differ in the number and types of inferences the readers make were revealed by Latent Profile Analyses (LPA). First, *Literal*

Readers stay close to the literal text; they predominantly repeat the text and engage in few inferential processes. Second, *Paraphrasing Readers* extract the meaning from the text; they predominantly paraphrase the text, but also make some text-connecting and elaborative inferences. Third, *Elaborating Readers* use background knowledge and go beyond as well as think ahead of the focal sentence; they predominantly use elaborative inferences, but also a fair number of text-connecting and predictive inferences.

2.4.1 Text Genres

To determine whether children approach various text genres differently, we examined think-aloud responses for both narrative and expository texts. For both types of texts, the profiles of Literal Readers, Paraphrasing Readers, and the Elaborating Readers emerged. There were considerable similarities in how the three subgroups approached narrative and expository texts. However, there also were a few noticeable differences. Children that differed in think-aloud response profiles across text genres tended to go beyond the text and use inferences when reading narratives but stay closer to the literal text and make fewer inferences when reading expository text. This resulted in more children belonging to the Literal and the Paraphrasing Readers in expository text reading than in narrative text reading. Furthermore, expository texts elicited fewer inferences within each profile. For example, there was a decrease in the number of predictive inferences made while reading expository texts compared to narratives. The decrease was particularly noticeable in Elaborating Readers (who produced most predictive inferences while reading narratives), but was also present in Paraphrasing and Literal Readers. Expository and narrative texts make different demands on the reader (e.g. Eason et al., 2012; Lorch, 2015). For example, narratives often include words that are part of everyday vocabulary whereas expository text may include more non-frequent technical terms (Medina & Pilonieta, 2006). In the current texts, a slight lower word frequency and a higher number of unique words were observed for expository texts compared to narratives. In addition, narratives activate familiar script structures (with initiating goals, actions, and outcomes) similar to everyday experiences (e.g., Britton & Pelligrini, 1990) that likely enhance predictive inferences (Narvaez, van den Broek, & Barrón Ruiz, 1999), whereas expository text structure may not offer the constraints necessary for making predictions (e.g., Cook et al., 2001; Klin et al., 1999). In the current study, the expository texts contained more connectives for higher text cohesion, however, the narratives were possibly experienced as easier to comprehend because of higher familiarity of word and text structure, as also argued by Graesser, McNamara, and Louwerse (2003). Our results indicate that differences in text demands elicit differences in children's reading behavior.

Expository texts did not only elicit a decrease in overall inference generation, but also an increase of invalid elaborative inferences. The subgroup producing most valid inferences - Elaborating Readers – also generated most invalid elaborative inferences, but there was also an increase in Paraphrasing Readers. Because invalid inferences are likely to impede comprehension (cf. McMaster et al., 2012), this observation constitutes an important issue for future studies. For example, investigating reading processes that occur after invalid inferences may shed light on the impact invalid inferences have on local and global text comprehension. Questions pertaining to whether and how the positive influence of making valid inferences outweighs the negative effect of making invalid inferences and, to the interplay of valid and invalid inferences as they occur while a text unfolds are highly relevant.

Considering reading profiles under different conditions provides insight on the stability of readers' approaches to text and their use of specific sets of reading processes. On the one hand, research on reading profiles suggests that children possess a certain set of skills, a reading trait. Longitudinal research adds evidence that such skills are relatively stable and predict future performance (e.g. Kendeou, van den Broek, White, & Lynch, 2009). On the other hand, research investigating different task- and text demands (for a recent review, Kendeou & van den Broek, 2017) suggests that reading processes may fluctuate depending on the situation, and that readers' standards of coherence moderate the processes in which they engage (e.g. van den Broek & Helder, 2017). The current findings show that children produce a very similar set of think-aloud responses across text genres but also adjust to the different situations. Hence, the results suggest a certain stability in children's text approaches with room for situational differences.

2.4.2 Reader Characteristics

Including readers from a large range of poor to good comprehension ability enabled us to investigate whether the children with different reading profiles differ in underlying competences. Based on narrative texts, Elaborators perform significantly better than the other two groups on word decoding, reading comprehension, and non-verbal reasoning. Based on expository texts, Elaborators perform significantly better than Literal Readers on reading comprehension and non-verbal reasoning. Thus, children with good comprehension and non-verbal reasoning abilities also make more frequent use of background knowledge and make more predictions about the upcoming text. This is in line with previous findings that reading skills (e.g., Cain & Oakhill, 1999; Carlson et al., 2014) and non-verbal reasoning skills (de Leeuw et al., 2016; Naglieri, 2001) are positively related to inference making. The causal relation between inference generation and reading comprehension may take different forms; inference generation may aid text comprehension, but also good word and text comprehension may provide possibilities to connect the text to background knowledge. Interestingly, current results suggest that non-verbal reasoning skills are also important in this equation. These findings encourage future research to investigate whether the relation between inference generation and comprehension is mediated by reasoning abilities.

The profiles did not differ significantly on vocabulary or working memory. However, and perhaps to be noted as a limitation of the study, it is difficult to dismiss these measures as unimportant when investigating individual differences. As can be seen in the supplementary Table 2.S2, all ancillary measures correlate with each other, with the exception of word decoding which did not correlate with vocabulary and reasoning skills. Furthermore, it is worth mentioning that, numerically, Elaborating Readers (based on both narrative and expository reading) performed better than readers with the two other profiles on all five measures of language abilities and cognitive resources. Similarly, Paraphrasing Readers numerically performed better than Literal Readers on all five measures. This consistency suggests a hierarchy between the three profiles in which better language abilities and cognitive resources are related to the ability to make inferences that go beyond the literal text.

2.4.3 Paraphrasing Text Facilitates Inference Generation

There were considerable differences in processing texts between Literal and Paraphrasing Readers. Not only do Paraphrasing Readers make more paraphrases and fewer repetitions than Literal Readers do, they also produce more inferences. In light of previous findings that paraphrasing a text may help readers start up inferences (e.g., Bohn-Gettler & Kendeou, 2014), one may speculate that the Paraphrasing Readers profited from being able to summarize the text in their own words to enable inference generation. Conversely, being less able to extract the meaning beyond the literal level of a text may hamper inference making and reading comprehension in Literal Readers.

2.4.4 Levels of Reading Comprehension

An interesting aspect of the observed three reading profiles is that they resemble ideas about different levels of text comprehension: surface, textbase, and situation-model understanding (e.g., van Dijk & Kintsch, 1983; Kintsch, 1988, 1994). Whereas Literal Readers mainly show a surface level understanding of the text with many verbatim repetitions, Paraphrasing Readers mainly show a textbase understanding by focusing on the *meaning* of the text input. Finally, Elaborating Readers create an enriched situation model of the text by drawing on background knowledge. Again, this suggests a hierarchy among the reader profiles, with Elaborating Readers being most likely to attain deep comprehension. Empirical support for such hierarchy comes from two of the current findings. First, Elaborating Readers, showed best comprehension and reasoning abilities, followed by the Paraphrasing Readers and, thereafter, the Literal Readers. Second, profile membership across text genre was rather stable, but children that differed in reading profile across text genres seemed to follow a hierarchical pattern: some individuals who were identified as Elaborating Readers while reading narrative texts were identified as Paraphrasing Readers while reading expository texts, but none changed his or her response pattern so drastically as to become a Literal Reader. Together, these findings suggest a hierarchy between the three profiles with reading comprehension ranging from basic to deep understanding. Whether these profiles are stable over time or reflect stages in children's reading development is an empirical question which can only be answered by longitudinal research.

2.4.5 Educational Implications

The central findings of the current research are that readers systematically differ in how they process text and that these differences are associated with inference generation and comprehension abilities. In an educational context, they suggest the importance of assessing the extent to which individual readers are able to enrich their mental representation of text with background knowledge (e.g., Carlson et al., 2014), and of using such assessment for individualized support of deep reading comprehension. For example, a paraphrasing training may benefit Literal Readers and encourage them to move beyond the literal meaning of the text to eventually enable inference generation (Bohn-Gettler & Kendeou, 2014; McNamara, 2004; McNamara et al., 2009). Similarly, questioning techniques have proven useful to encourage readers to generate more inferences (e.g., McMaster et al, 2012), which may be particularly beneficial for Paraphrasing Readers. Elaborating Readers use inferences correctly for narrative texts and displayed best performance on an ancillary comprehension measure. However, they make more invalid inferences for expository than for narrative texts. For these students it may be useful to focus on awareness and acknowledgement of invalid inferences (e.g., Van den Broek & Kendeou, 2008) and teach strategies that may compensate for lack of background knowledge. Teaching strategies for comprehension monitoring and metacognitive processes (e.g. McCrudden & Kendeou, 2014) could increase awareness of whether one's background knowledge matches the topic of the text.

In addition, reading methods such as the Concept Oriented Reading Instruction (CORI; e.g. Guthrie & Alao, 1997) may enhance motivation, increase strategy use, and increase conceptual knowledge during reading. Hence, we encourage tailoring instruction to individual needs, however with the aim of stimulating higher-order thinking and deep comprehension for all types of readers (Mayer, 1996; Oakes, 2008).

2.4.6 Reading Profiles; Additions and Limitations

Research on reading profiles with person-centred analyses (e.g. LPA) is promising as it addresses the fact that not all measures of reading processes are linearly related but are heterogeneous both between and within individuals. For example, the Literal and Elaborating Readers display almost opposite patterns in their think-aloud responses; Literal Readers make many text repetitions and few inferences, and Elaborating Readers make few text repetitions and many inferences. In addition, a reading profile between these two emerges, the Paraphrasing Readers. These differences may not have been picked up on using variable-centred analyses (e.g. linear models) which assume relations to be the same for all individuals. In this respect, current results support and expand previous literature on reading profiles. By distinguishing between close-to-text processes (text repetitions and paraphrases), it appears there are at least three reading profiles rather than the two reported before (e.g. Carlson et al., 2014; McMaster et al., 2012; Seipel et al., 2017); in addition, the profiles apply to a wider population than just poor or just good readers. However, the limited sample size of the current study did not allow to replicate the LPAs with a split-sample method. Hence, we could not test the generalizability of these three reading profiles, neither in a sample with a range of reading proficiency, nor at different levels of reading proficiency. Further research alternating the pool of participants and the reading task is still needed to generalize reading profiles.

Our results were obtained with a think-aloud method and we make interpretations within the possibilities of this method. However, we believe it is important to discuss some limitations of the think-aloud method. First, thinking aloud while reading may focus a reader's attention on processes he or she is not normally aware of and, thereby, alter the reading experience and text comprehension. However, given that in the current study the reading profiles are fairly stable across text genres and that Elaborating Readers systematically score highest on reading comprehension and reasoning abilities, evidence suggest that the obtained profiles are not just an artefact of the think-aloud method. Second, developing readers may not have enough attentional capacity to extensively reflect on all higher-order cognitive processes occurring while reading. If individual differences in reporting far-from-text processes (i.e. inferences) are due to an inability to formulate thoughts on ones' reading processes, then distinguishing between two kinds of close-to-text processes (i.e. text repetitions and paraphrases) takes on added importance, as it allowed investigation of qualitative differences between young readers' processing of the literal text.

2.5. Conclusions

We identified three reading profiles -Literal, Paraphrasing, and Elaborating Readers- that differ in the number and types of inferences they make. The reading profiles were remarkably constant across narrative and expository texts. However, for expository texts children tend to fall back on reading processes closer to the literal text. Furthermore, differences between these profiles reveal a positive relation between the ability to elaborate on the text at hand and comprehension ability as well as reasoning ability.

The identification of profiles has implications for reading theory as well as for educational practice. On the theoretical side, it deepens our understanding of individual differences in inference generation in particular, and of the complex interaction of processes involved in reading comprehension in general. On the educational side, it provides a basis for conceptualizing individualized and adaptive approaches to reading instructions with the aim to promote in-depth comprehension.

Supplementary
Think-Aloud Texts Translated from Dutch to English

Mieke Picks a Present

One day, Mieke remembered that her mother's birthday was coming soon. She wanted to give her mother a nice present. Mieke went to the store and found a nice pair of earrings. She bought them for her mother, went home, and wrapped them in bright paper. When the present was nicely wrapped, she hid it in her closet.

The next day, Mieke saw her friend Sandra make her own jewellery. When she was little, she had made a bracelet and she remembered that her mother liked hand-made jewellery. She decided to make a nice necklace for her mother. She chose the pearls she wanted to use and followed Sandra's instructions. Finally, Mieke had made a beautiful and long necklace. Mieke made a knot and tied a hook to the string. She stored it in a beautiful red box in her closet until the big day. Mieke was very happy with her decision to make a necklace and returned the earrings to the store.

At last, her mother's birthday arrived. Mieke took the red box from her closet. Her mother opened the box and was very happy when she saw the necklace. She thanked Mieke for the beautifully hand-made gift.

When the Earth Shakes

Sometimes earthquakes happen in South of Europe. An earthquake occurs when the ground suddenly begins shaking. About 100 kilometres inside the earth are layers of rock that move and crack. It is the moving and cracking of these rock layers that causes earthquakes. Each year, many small earthquakes affect the South of Europe. Even though most people normally can't even feel small earthquakes, large earthquakes can cause major damage. They can cause buildings, bridges, and homes to fall down.

Because earthquakes are so common, buildings need to be built in special ways. Although this helps protect buildings from earthquakes, damage is still possible. Recently, large earthquakes have shaken several cities in South of Europe. For example, one woman recollected a memory for us about a large earthquake that happened in Rome in 2009. First, she heard a loud crack. Then, the dining area of her second-floor apartment moved downward. "I felt like I was falling," she recalled. "Until I was outside of the building, I was unsure of what had happened. The earthquake had completely crushed the first floor."

Holiday Shopping

Lisa thought it was time to go shopping for Holiday presents. She disliked shopping when the stores were busy. When she got to the mall it was crowded, the parking lot was full.

She was determined to buy her three-year-old daughter an adorable doll. She found one, paid for it, and left to continue shopping for other things.

Lisa had no idea what to get her husband. She went to a men's clothing shop and looked at expensive ties, suits, and cologne. Then she went to a tech store and examined laptops, MP3-players, and televisions. Nothing at these places seemed appropriate. She walked by a pet store and saw an adorable puppy. She knew instantly that this was the perfect gift. She left the mall and decided to postpone more shopping until the next day.

Lisa got in her car and drove home. She realized she couldn't bring the puppy home. She went to a friend's place and asked if she could leave the puppy there. Her friend said yes and invited her inside for some thee.

Dutch Snakes

In the Netherlands the temperature is most often agreeable for humans. This is not the case for snakes. They don't think it is hot enough. Because of this there are only three types of snakes in the Netherlands, but they are rarely sighted.

A snake that we do see more often is the collar snake, this snake is active during the day. Collar snakes can be recognized by the yellow collar behind their head. Their food consists of frogs, lizards, and fish, and they can swim very well.

The viper is the only poisonous snake in the Netherlands. You may encounter this snake at the Veluwe or in Drenthe. The poison is deadly for their normal pray: field mice and lizards. Vipers live all over Europe, even in North of Russia. Because of the cold there it takes two years for the eggs to hatch.

In the Netherlands there is also a small constrictor, this snake is not poisonous. They can become up to 75 centimetres long and eat preliminary lizards and mice. You can recognize constrictors by the black and brown stripe on both sides of their heads. They mostly live in dry areas such as moors, grass fields and at the outskirts of forests.

Supplementary Descriptive Analysis

Table 2.S1

Means, Range, Standard Deviation, Skewness and Kurtosis for All Measures are Reported here for All Children. The (N) and the (E) Signals Descriptive Statistics for the Narrative and the Expository Texts, Respectively.

	Mean	Range	Standard Deviation	Skewness	Kurtosis
Word Decoding	97.22	44.00	9.83	.86	.60
Reading Comprehension	41.18	87.00	18.44	.83	.18
Reasoning Ability	38.43	40.00	8.04	-.71	1.18
Vocabulary	41.63	33.00	6.58	-.16	-.26
Working Memory	8.40	21.00	4.68	.48	-.16
Text Repetition (N)	27.80	88.57	22.36	.64	-.42
Paraphrase (N)	22.08	51.43	12.04	.29	-.50
Text-connecting Inference (N)	11.80	42.86	8.37	1.10	1.43
Valid Elaborative Inference (N)	18.29	65.71	14.26	.85	.23
Invalid Elaborative Inference (N)	1.33	14.28	2.52	2.53	7.98
Predictive Inference (N)	9.67	57.14	11.47	1.88	4.04
Text Repetition (E)	29.71	84.85	23.46	.58	-.63
Paraphrase (E)	23.11	60.61	11.86	.25	.20
Text-connecting Inference (E)	9.63	33.33	6.30	.88	.97
Valid Elaborative Inference (E)	18.41	57.57	14.20	.71	-.23
Invalid Elaborative Inference (E)	2.80	15.15	3.50	1.61	2.63
Predictive Inference (E)	2.55	21.21	4.06	2.12	5.15

Supplementary Correlational Analysis

Table 2.S2

Correlations among Word Decoding, Reading Comprehension, Reasoning Ability, Vocabulary, and Working Memory are Reported here for All Children (N=104).

	1	2	3	4	5
1. Word Decoding	1	-	-	-	-
2. Reading Comprehension	.358**	1	-	-	-
3. Reasoning Ability	.108	.530**	1	-	-
4. Vocabulary	.179	.618**	.459**	1	-
5. Working Memory	.194*	.436**	.340**	.272**	1

Note. Asterisks indicate significant levels at * $p < .05$, ** $p < .01$.

Supplementary Post Hoc ANOVAs

Table 2.S3

Posthoc ANOVA comparisons between the subgroups. In addition to the LPA's visualized in figure 2.1 and 2.2, we here show the p-value of the Posthoc ANOVAs to display whether the subgroups differ significantly between each other on each think-aloud response category. LR. = Literal Readers, PR = Paraphrasing Readers, and ER = Elaborating Readers. The (N) and the (E) Signals Statistics for the subgroups based on Narrative and Expository Texts, Respectively.

	LR vs. PR (N)	LR vs. ER (N)	PR vs. ER (N)	LR vs. PR (E)	LR vs. ER (E)	PR vs. ER (E)
Text Repetition	.000	.000	.000	.000	.000	.000
Paraphrase	.000	.008	.000	.000	.781	.000
Text-connecting Inference	.001	.000	.052	.010	.301	.432
Valid Elaborative Inference	.000	.000	.000	.000	.000	.000
Invalid Elaborative Inference	-	-	-	.043	.000	.000
Predictive Inference	.368	.000	.000	.940	.063	.013

Chapter 3

The Effect of Upper Elementary School Children's Online Reading Profiles on their Memory Representation of Narrative and Expository Texts

Karlsson, J., van Leijenhorst, L., & van den Broek, P. (in preparation). The effect of upper elementary school children's online reading profiles on their memory representation of narrative and expository texts.

The Effect of Upper Elementary School Children's Online Reading Profiles on their Memory Representation of Narrative and Expository Texts

Abstract

We examined 9-11-year-old children's ability to selectively recall more central information than peripheral information from text (the recall centrality effect) as a function of their online reading profiles (Elaborating, Paraphrasing, and Literal Readers) and text genre (narrative and expository texts). Elaborating Readers generate more inferences than Paraphrasing Readers, who in turn generate more inferences than Literal Readers. We anticipated that children in profiles that generate a larger number of inferences while reading would show a larger centrality effect because making inferences help in understanding what is interconnected, and hence central, in a text. Children in all reading profiles showed a centrality effect for narrative text. Elaborating Readers showed a larger centrality effect than Paraphrasing Readers, but Literal Readers perform similar to both other profiles. There was no centrality effect for expository texts for children in any of the three reading profiles. We conclude that the use of online inferences alone does not predict differences in text memory and suggest that children in the reading profiles use the text approach that is most suitable for their cognitive abilities to understand the text.

Keywords: Reading Profiles, Online Processes, Centrality Effect, Offline Representation, Text Genres

3.1 Introduction

To understand a text, readers engage in various online reading processes to form a structured and coherent offline memory representation of that text, a situation model (van Dijk & Kintsch, 1983). This memory representation enables readers to remember the meaning of the text, and to learn from it. Children need to build an extensive repertoire of reading processes and learn to apply them optimally, eventually more or less automatically, to different types of texts. Children in elementary school differ in the numbers and types of online reading processes they engage in while reading. Recent studies show that developing readers can be classified into distinct reading profiles based on their online reading processes (Carlson, Seipel, & McMaster, 2014; Karlsson et al., 2018 (chapter two); Kraal, Koornneef, Saab, & van den Broek, 2017; McMaster et al., 2012). These profiles suggest that children use different approaches to text: some stay close to the literal meaning of sentences whereas others generate more inferences to make connections between sentences and with their background knowledge. This research has given us insight into individual differences in the collective set of online reading processes in which beginning readers engage. However, much less is known about possible differences in the memory representation that result from these profiles' online reading processes. Because the memory representation is the basis for learning from texts in school, it is important to understand if children with different reading profiles create different memory representations.

Beginning readers initially learn to read in the context of narrative texts, later they are required to read expository texts that are used for knowledge transfer in higher grades. Because the two text genres differ in their demands on the reader and require partly different reading processes (Lorch, 2017), beginning readers find narrative texts easier to comprehend than expository texts (e.g. Best, Floyd, & McNamara, 2008). Therefore, to gain a more wholistic understanding of how children's online reading processes affect their resulting memory representation, it is important to consider the influence of text genre. In the current study we investigate the effect of children's (9-11 years old) online reading profiles on their memory for narrative and expository text. We do so by examining their ability to build and recall a memory representation, structured around the central information in the text.

3.1.1 Online Text Processes and Offline Text Representation

A general consensus has emerged from scientific models of reading comprehension that readers'online reading processes affect their offline text representation (McNamara & Magliano, 2009; van den Broek, Young, Tzeng, & Linderholm, 1999; van Dijk & Kintsch, 1983). The text representation that readers build can vary in depth and richness of information. For example, readers' text comprehension ranges from a surface level, to a literal understanding, to a situation-model level. The latter is an elaborate mental representation where information from the text is combined with background knowledge to yield an in-depth understanding of the meaning of the text (van Dijk & Kintsch, 1983). These qualitatively different representations are the result of the cognitive processes in which readers engage during reading. Use of elaborate online processes is thought to lead to a more elaborate mental representation. Rereading parts of the text is an example of relatively less elaborate processing, whereas generating inferences using different types of background knowledge is an example of more elaborate processing (e.g. Karlsson et al., 2018; Magliano & Millis, 2003; Seipel, Carlson, & Clinton, 2017). Inferences in particular are important for building a situation model because they help readers create connections between units of information within the text and interrelate these connections with their background knowledge (Graesser, Singer, & Trabasso, 1994; O'Brien, Cook, & Lorch, 2015; van den Broek, 1990). Inference generation has been positively related to good text comprehension in both adult (Graesser et al., 1994; Tun, 1989) and developing readers (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Lynch et al., 2008; McGee & Johnson, 2003).

A good mental representation of a text not only contains the information that the text conveyed, but also differentiates information that is central in the text. Information that is central describes the gist or main idea of the text. Central information is usually strongly connected within the text (e.g. Omanson, 1982a; van den Broek, 1990) and, thus, forms an interconnected set of concepts that provide strong memory access routes for readers when they try to recall the text (Albrecht & O'Brien, 1991; O'Brien, & Myers, 1987). Therefore, the text representation of reading-proficient adults and children shows a *centrality effect*, i.e. readers show better memory for central than for peripheral text information (Albrecht & O'Brien, 1991; Armbruster, Anderson, & Ostertag, 1987; Bauer & San Souci, 2010; Kendeou, van den Broek, White, & Lynch, 2007; Mo, Chen, Li, Chen, & He, 2007; Miller & Keenan, 2009; Yeari, Oudega, & van den Broek, 2016; Yeari, van den Broek, & Oudega, 2015; Yekovich & Walker, 1986). To summarize, a good memory representation of a text is the result of cognitive processes during reading and is sensitive to the centrality of information from the text. An open question is if individual differences in online comprehension processes, such as those captured by the different reading profiles, translates into offline memory representations that systematically differ with respect to the centrality effect.

3.1.2 Reading Profiles

Investigations of online reading processes often focus on a particular process, such as inference generation, in isolation of other processes. However, readers' collective set of processes while reading a full text are equally important to consider, to understand how each process affects the next (van den Broek et al., 1999) and how readers gradually form an offline representation of the text they are reading (Hyönä, Lorch, & Kaakinen, 2002). Cluster analyses of children's overt thoughts while reading (think-aloud responses) have revealed distinct reading profiles in elementary school children. Although children engage in a variety of processes while reading, they differ in the kinds of processes in which they engage the most (Carlson et al., 2014; Karlsson et al., 2018; Kraal et al., 2017; McMaster et al, 2012; Seipel et al., 2017). Children who go beyond the literal text and extensively generate inferences have been called *Elaborators* (Carlson et al., 2014; Karlsson et al., 2018; Kraal et al, 2017; McMaster et al, 2012; Seipel et al., 2017). Elaborating readers predominantly use text-connecting, elaborating, and predictive inferences to connect the information in the text to their existing knowledge of the world. Children who stay close to the literal text make up other reading profiles, such as *Paraphrasers.* In previous research, Paraphrasers respond with many literal repetitions and paraphrases (Carlson et al., 2014; Kraal et al, 2017; McMaster et al, 2012; Seipel et al., 2017). However, we recently found a further refinement as cluster analyses also revealed the profile *Literal Readers* (Karlsson et al., 2018 (chapter two)). In this study Paraphrasing Readers predominantly rephrase the text into own words but also generate some inferences, whereas Literal Readers predominantly repeat the text verbatim (Karlsson et al., 2018).

Using background knowledge to generate inferences while reading has a positive impact on the offline mental representation (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Lynch et al., 2008; McGee & Johnson, 2003). Specifically, inferences that recruit relevant background knowledge help readers make connections between text events and recognize central aspects of the text (van den Broek, 1990). Because children in the three reading profiles (Elaborating, Paraphrasing, and Literal Readers) differ in the number of such inferences they generate, they may show differences in their ability to extract central text information. Therefore, we expect children in profiles characterized by more inference generation to show a stronger centrality effect than children in profiles characterized by less inference generation.

3.1.3 Text Genre Differences

Text genre influences reading processes and memory of text content (Best et al., 2008; McNamara, Ozuru, & Floyd, 2011; Tun, 1989; Wolfe, 2005; Wolfe & Woodwyk, 2010; Zabrucky & Ratner, 1992). Two genres are frequently contrasted in reading research as they differ from one another in multiple ways: narrative and expository texts. Narrative text intends to tell a story in which the goal of a protagonist and the ensuing actions to fulfil that goal form a causal/temporal line, from beginning to middle and end (Mandler & Johnson, 1977; Omanson, 1982a; Wolfe, 2005). Narratives cover events that are often similar to readers' everyday experiences and use everyday vocabulary (Britton & Pelligrini, 1990; Medina & Pilonieta, 2006). The causal story line and well-known words ease comprehension processes: readers fairly easily can read a sentence, make associations and inferences to relevant concepts of previous text or background knowledge, and then evaluate and integrate these pieces of information (Lorch, 2017; Omanson, 1982a; van den Broek, 1990). Expository texts intend to convey information to the reader. They tend to be more descriptive, often explaining a main idea and subordinate facts but do not necessarily entail a causal/temporal line that facilitates readers' memory for the textual information (Meyer, 1985; Taylor, 1980; Tun, 1989; Wolfe, 2005). In addition, expository texts often use novel words and concepts, abstract relations, and a higher information density compared to narrative texts. These properties tend to increase processing demands on readers and reduce inference making (Best et al., 2008; Lorch, 2017; Zabrucky & Ratner, 1992). Indeed, children in all reading profiles make fewer inferences when reading expository texts than when reading narrative texts (Karlsson et al., 2018). Therefore, we expect children to show a stronger centrality effect for narrative texts than for expository texts.

3.1.4 Current study

This study is part of a larger project. In the first part of the project we identified three reading profiles in 9-11-year-old children based on their online reading processes, think-aloud responses to text (Karlsson et al., 2018). In the current chapter, readers in the profiles Elaborating, Paraphrasing, and Literal Readers are compared on their ability to form a structured offline memory representation of narrative and expository texts, i.e. recalling information that is central to each text. Following prior research (Albrecht & O'Brien, 1991; Miller & Keenan, 2009; 2011; Omanson, 1982a, 1982b), we determined centrality of text units by importance ratings. We then examined the children's recall of units rated as highly important and central to the text, and units rated as peripheral and less important.

There are two research questions. First, do the children with different reading profiles show similarities or differences regarding the centrality effect? We anticipate Elaborating Readers, who routinely use background knowledge and generate inferences, to show a larger centrality effect compared to Paraphrasing and Literal Readers. This would be reflected in a larger difference in their recall of central and peripheral text units. Furthermore, Paraphrasing Readers engage in more inference generation than Literal Readers and are expected to show a larger centrality effect than Literal Readers. Second, do any observed effects depend on text genre? Children in all profiles generated fewer inferences and showed an approach closer to the text for expository than for narrative texts (Karlsson et al., 2018). Therefore, it is possible that readers in all profiles show a larger centrality effect in their recall of narrative texts than of expository texts.

3.2 Methods

3.2.1 Participants

One hundred and seven children (61 girls) between nine and eleven years old (M = 10.3, SD = 0.73) participated in the current study. Children from grades six and seven were recruited from 16 schools in the south-west of the Netherlands. Inclusion criteria were typical development, Dutch as native language, and average to good word decoding skills (the latter meaning scoring above the 40th percentile on the Dutch national standardized test, CITO Drie Minuten Toets, DMT; Krom, Jongen, Verhelst, Kamphuis & Kleintjes, 2010). Parents gave written consent and children gave oral consent for participation. The study was carried out in accordance with the Declaration of Helsinki and was approved by the Ethical Committee at Leiden University.

3.2.1.1 Defining Reading Profiles

In this study we analyse the recall data of participants from three different reading profiles that were identified in the first part of this project (for a detailed description of the procedure see Karlsson et al., 2018). The reading profiles were characterized by their think-aloud responses while reading. First, the think-aloud responses from each child were assigned to six categories. Second, the response data were used to create six variables: the percentage of text repetitions, paraphrases, text-connecting inferences, valid elaborative inferences, predictive inferences, and invalid elaborative inferences. These steps were conducted for the narrative and the expository texts separately. Third, two Latent Profile Analyses (LPA; Hagenaars & McCutcheon, 2002; Oberski, 2016) were performed on the responses from reading narrative and expository texts, respectively.

The Latent Profile Analyses (LPA) revealed three distinct reading profiles that differ in the number and types of inferences the readers make (Karlsson et al., 2018). The same reading profiles were found for both narrative and expository texts. For narratives, the profiles Literal, Paraphrasing, and Elaborating Readers comprised 23, 46, and 37 children, respectively. For expository texts, the profiles Literal, Paraphrasing, and Elaborating Readers comprised 29, 50, and 27 children, respectively. A significant majority, 75.4%, of the children had the same profile membership for narrative and expository texts (Karlsson et al., 2018). Elaborating Readers mainly use elaborative inferences, explaining the text to themselves, but also use text-connecting- and predictive inferences. Paraphrasing Readers mainly extract the meaning from the text, paraphrasing, but also engage in some text-connecting and elaborative inferences. Literal Readers focus on repeating the text verbatim during the think-aloud task and engage in few inferential processes. Expository texts generally elicited a smaller number of inferences. In the current paper, two mixed design ANOVAs are used to analyse children's recall. For these analyses, the children's profile membership for narrative and for expository texts, respectively, will be used as between-subject variable.

3.2.2 Texts and Tasks

Participants read and were asked to recall two narrative and two expository texts. The narrative texts followed a typical story structure (e.g., Mandler & Johnson, 1977). The narratives began by introducing a protagonist who had a specific goal, which for both stories was to find a suitable gift for a family member. The narratives then sketched the protagonist's attempts to fulfil this goal and ended with the successful attainment of the goal. The expository texts followed a descriptive text structure with main and subordinate ideas (Meyer, 1987). The texts included different kinds of elements such as causal relations and descriptions of places. One text about snakes in the Netherlands explained the effect of temperature on hatching eggs and described well-known places where snakes live. The second text explained the cause of earthquakes and locations in Europe where they commonly occur. Texts from the two genres were fairly similar with respect to text-length characteristics (measured with P-CLIB version 3.0; Evers, 2008), and syntactic complexity (measured by the Development Level with T-Scan; Pander Maat, Kraf, & Dekker, 2017). One of the expository texts had a higher grade difficulty, because of a lower percentage of frequent words. In addition, the expository texts had a higher information density than narratives as indicated by the type-token ratio. However, they also showed higher text cohesion as indicated by a higher density of connectives. Further data on text characteristics and English translations of the four texts are published in Karlsson et al., 2018.

Each child received a booklet with the four texts, with each sentence printed on a separate page. Participants were instructed to read each sentence aloud and then report what they were thinking before continuing with the next sentence (for a detailed description of the procedure, see Karlsson et al., 2018). At the end of each text participants were asked to recall the text. To encourage the children to recall as much as possible they were asked to retell the whole text, as if they were telling the story to a friend who had not heard it. Children were prompted twice at the end of each recall with the question 'Is there something else you can remember?'

3.2.2.1 Defining Central and Peripheral Text Units

To distinguish between central and peripheral text units, each text was rated by 30 native Dutch speaking undergraduate students (15 women, mean age 20 years). The students provided informed written consent and were then tested individually at the university. They received a small payment or course credits for participation. Each text was parsed into subject-verb clauses (idea units), yielding a total of 56 and 40 idea units for the two narrative texts and 47 and 35 idea units for the two expository texts, respectively. To rate the texts the students received a booklet with a paper-and-pencil task. Each text was presented as a regular text on one page. A list of the idea units of that text was presented on the next page. After having read the text, students were asked to rate on the list version how important each unit was to understand the text as a whole. Importance was rated on a scale from 1-7, with 1 being least important and 7 most important. There was a high agreement among students' rating on all four texts, all Cronbach alpha coefficients > .93, and all ICC (2, 30) > .935 (Landers, 2015). The units for each text were divided into three equal parts: top, middle, and bottom ranked units. The top third and bottom third ranked units were used to examine children's memory of central and peripheral units, respectively. The respective mean rating for top and bottom thirds were 5.05 ($SD = 0.63$) and 2.22 ($SD = 0.65$) for the narrative texts and 4.98 ($SD = 0.67$) and 2.80 ($SD = 0.66$) for the expository texts. A repeated measures ANOVA showed an interaction effect of text genre by importance level, $F\ (1, 29) = 30.04$, $p < .001$, $\eta_p^2 = .51$. Follow-up analyses showed that top ratings were significantly higher than bottom ratings for both narrative ($F\ (1, 29) = 412.66$, $p < .001$, $\eta_p^2 = .93$) and expository texts ($F\ (1, 29) = 223.91$, $p < .001$, $\eta_p^2 = .88$). Whereas top ratings did not differ between the two text genres ($F\ (1, 29) = 0.71$, $p = .408$, $\eta_p^2 = .02$), bottom ratings for expository texts were higher than those for narrative texts ($F\ (1, 29) = 38.17$, $p < .001$, $\eta_p^2 = .57$).

3.2.3 Procedure

The recall data for the children was gathered in a quiet room at school in an individual session. During this session children received the combined think-aloud and recall task, as well as additional tasks described in the previous publication (Karlsson et al., 2018). The session lasted approximately one hour, including instructions and a short break. Participants were given verbal instructions and one practice text. After completion of the experiment children received a small gift to thank them for their participation.

The think-aloud and recall session for each participant was recorded and later transcribed. Two test leaders categorized the recall data with good interrater reliability, $K = 0.76$, $p < .001$. One point was given for each accurately remembered idea unit. For each participant, averaged recall scores were computed for central and peripheral units, respectively. This was done separately for the narrative texts and the expository texts. This created four within-subject conditions: recall of central and peripheral units in narrative and in expository texts, respectively.

3.2.4 Analyses

We analyzed children's recall of central and peripheral text units as a function of their reading profiles. Two separate mixed-design analyses of variances were performed, one for narrative and one for expository texts. No outliers were found. Recall data from one child were partly missing due to technical problems, therefore, the analyses include data from 106 children. In the first 2 (recall of central and peripheral narrative text units) x 3 (reading profiles) mixed design ANOVA, profiles were based on narrative texts. In the second 2 (recall of central and peripheral expository text units) x 3 (reading profiles) mixed design ANOVA, profiles were based on expository texts.

3.3 Results

3.3.1 Recall of Narrative Text

In the 2 x 3 ANOVA, a main effect of centrality revealed better recall of central than of peripheral units, $F\ (1, 103) = 457.02$, $p < .001$, $\eta p^2 = .82$ (see Figure 3.1). There was no main effect of reading profile, $F\ (2, 103) = 1.40$, $p = .253$. However, there was a centrality-by-profile interaction, $F\ (2, 103) = 4.30$, $p = .016$, $\eta p^2 = .08$. This interaction was examined in three steps. First, separate follow-up ANOVAs for each of the three reading profiles showed that children in all profiles recalled more central information than peripheral information, all ps $< .001$ (see Figure 3.1a). Second, we calculated a difference score for each child by subtracting the percentage of recalled peripheral units from the percentage of recalled central units. An ANOVA comparing the reading profiles on the children's difference scores showed that the Elaborating Readers had a larger difference score (*M* = 29.99, *SE* = 1.95) than did Paraphrasing Readers (*M* = 22.37, *SE* = 1.75), $p = .013$ (see Figure 3.1a). Literal Readers (*M* = 24.68, *SE* = 2.47) did not differ from Elaborating or Paraphrasing Readers, *p*s > .285. Third, follow-up ANOVAs examining central and peripheral units separately, revealed no significant differences in the number of central units recalled across the profiles ($p > .353$) and no differences in the number of peripheral units recalled ($p > .162$).

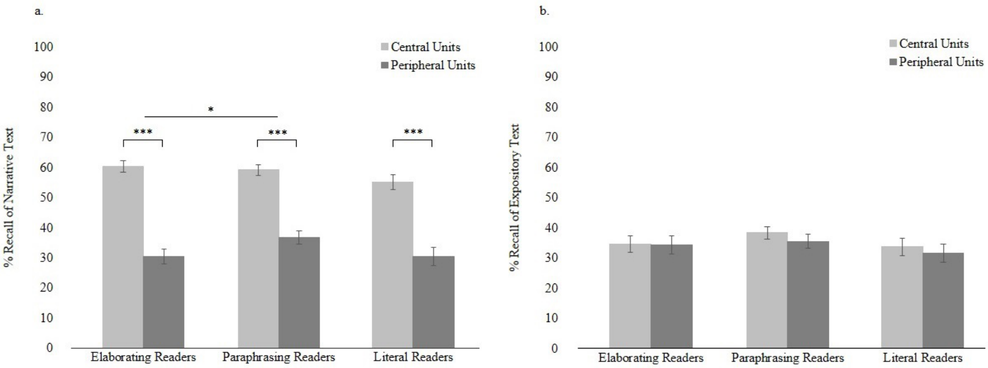

Figure 3.1. Mean percentage recall of central and peripheral units in narrative text (a) and expository text (b) by the three reading profiles, Elaborating, Paraphrasing, and Literal Readers. For narratives, children in all profiles showed a centrality effect of $p < .001$ (***), also, Elaborating Readers showed a larger centrality difference score than Paraphrasing Readers, $p = .013$ (*).

3.3.2 Recall of Expository Text

The 2 x 3 ANOVA, revealed no significant main effect of centrality, $F(1, 103) = 1.29$, $p = .259$ (see Figure 3.1b), no significant main effect of reading profile, $F(2, 103) = 0.95$, $p = .391$, and no centrality-by-profile interaction, $F(2, 103) = 0.26$, $p = .773$.

3.4 Discussion

The aim of this study was to examine whether the memory representation of narrative and expository texts of Dutch elementary school children (9-11 years old) differs as a function of their reading profile: Elaborating, Literal, and Paraphrasing Readers. Children in all three profiles showed a centrality effect in their recall of narrative texts. On average, they all arrive at a situation model in which central information is more prominent than peripheral information, a centrality effect. However, the Elaborating Readers distinguish between central and peripheral information in narrative texts more strongly than Paraphrasing Readers. Children's recall of expository texts did not show a centrality effect in any of the profiles, suggesting that all groups of children struggled to distinguish central and peripheral information in the expository texts.

3.4.1 Reading Profiles

In line with previous research (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Graesser et al., 1994; Kraal et al., 2017; Lynch et al., 2008; McGee & Johnson, 2003; van den Broek & Kendeou, 2017), we find that inference-making contributes to a readers' understanding of the text. Although children in the three profiles differ in the number of inferences they make, they all engage in inference generation during reading to some degree. Our findings suggest that this inference generation is sufficient to extract central information in narratives. However, the number of online inferences alone cannot fully predict the strength of the centrality effect in the resulting memory representation. Consistent with our predictions, the Elaborating Readers displayed a larger centrality effect than Paraphrasing Readers. However, counter to our predictions, memory for the text of Elaborating Readers and Paraphrasing Readers did not differ significantly from that of Literal Readers.

To better understand these seemingly contradicting results we may turn first to studies considering other reader characteristics and second to further reader characteristics of the children in the current three reading profiles. Previous research shows that seven-to-ten year old children develop various memory strategies; from strategies that are less cognitively demanding, such as repetition, to strategies that are more demanding, such as sorting, chunking, drawing on background knowledge, or combinations of strategies (Schneider, Kron-Sperl, & Hünnerkopf, 2009). Repeating content is more effective for text memory than using no strategies (Jonsson, Wiklund-Hörnqvist, Nyroos, & Börjesson, 2014), but the use and efficiency of strategies seems to vary with children's cognitive resources. For example, children with high Working Memory Capacity (WMC) benefit from using elaborate memory strategies, whereas children with low WMC benefit from using less elaborate strategies, such as repetition (Jonsson et al., 2014; Turley-Ames & Whitfield, 2003). In the previous paper (Karlsson et al., 2018), our Elaborating Readers proved not only to generate more inferences than the other two profiles while reading, but also showed better word decoding and non-verbal reasoning abilities in ancillary measures. With this knowledge, the results of this study may indicate that Elaborating Readers, who have good reading fluency and good reasoning abilities, utilize their resources to engage in elaborative processes that result in a well-structured memory representation of the text. Paraphrasing and Literal Readers with relatively poor reading fluency and reasoning abilities have fewer resources available for cognitively demanding reading processes compared to Elaborating Readers and pay more attention to the focal sentence, but are able to arrive at a structured text memory in this recall task. Similarly to previous research (Jonsson et al., 2014; Turley-Ames & Whitfield, 2003), we suggest that these three reading profiles use the text approach that is most suitable for their cognitive abilities to pursue the aim of comprehending the text. However, although Paraphrasers have comparably fewer cognitive resources at their disposal, they are perhaps trying out or developing more elaborate online reading strategies. Taxing their cognitive capacities in this way may, in comparison to Elaborating Readers, render a somewhat less structured memory representation, shown by a smaller centrality effect.

3.4.2 Text Genre Differences

The current results show that expository texts are challenging to recall for elementary school children. Compared to narrative texts, children from all profiles tended to recall less information from the expository texts as they struggled to extract central information. In the current study, 'centrality' of textual information was determined through ratings by college students. It is possible that children struggle with exposition in elementary school years as they still need to learn to recognize different text structures and build up a body of background knowledge to have reference points to new knowledge in exposition (Best et al., 2008; Meyer, Brandt, & Bluth, 1980; Taylor, 1980). The difficulties in recalling expository texts may have limited our ability to reliably specify differences in memory for central and peripheral information, as well as differences between profiles. However, recall of peripheral information is similar across text genres, suggesting that readers in all groups had at least a basic understanding of all texts, and their difficulties in recognizing central information must have a different cause. An alternative explanation may be that the results reflect differences between narrative and expository texts regarding the role that inferences play in centrality. For example, centrality in narratives may depend on numbers of connections and of inferences that create these connections. However, centrality in expository texts may depend more on logical structure (Lorch, 2017), such as process inferences that specify the detailed steps or dynamic characteristics of an event as it unfolds (Snow & RAND, 2002). Probing children's reading processes that contribute to logical text structures in expository texts may be a promising venue for further research on how children create a coherent mental model of expository text.

3.4.3. Limitations and Future Research

Pronounced profile differences in the think-aloud task but small profile differences in the recall task may stem from task differences. The think-aloud task may be a particularly sensitive tool to show children's processing differences in all types of readers; those who are still developing, and those who have already mastered the ability to generate inferences. Although the free recall scores do not show a ceiling effect, hence indicating that the task was not too easy for the children, it may have been less sensitive in picking up differences specifically related to inference generation in online text processing. For example, in contrast to a summarizing task, which demands the reader to extract the meaning of the text (e.g. Brown & Day, 1983), the free recall task encourages little individual interpretation and allows for a more general memory recollection such that children with varying processing abilities may show similar performance. Research linking online and offline reading tasks often shows that there is a relation between the two; however, results are often described as giving "tenuous" support that the online and offline measures under investigation really are tapping into direct comparisons of comprehension (Anmarkrud, McCrudden, Bråten, & Strømsø, 2013; Rapp & Mensink, 2011). Perhaps because online measures tap into moment-by-moment reading comprehension, whereas offline tasks measure post-reading comprehension, they will at best match only in part. Hence, showing exactly how online and offline comprehension processes are linked is difficult, because they are in part per definition different. The current results provide a similar picture; results in the offline task cannot be explained by performance in the online task alone, other reader characteristics seem just as important to consider.

Future endeavours of research into reading profiles may also investigate how comprehension differences transfer to long-term learning. Long-term memory of text, or learning over time, builds on online process in a similar way as offline comprehension (e.g. Beker, Jolles & van den Broek, 2017). However, memory decays over time. Examining long-term memory of text in these profiles may be informative in showing how the collective set of reading processes predict learning from text.

3.5 Conclusions

Because text memory forms the basis for learning in school, it is important to understand whether children with different reading profiles create different memory representations. The results of the current study show that children in the reading profiles Literal, Paraphrasing, and Elaborating Readers all build a structured memory of narrative texts. Subtle differences in children's narrative text memory could not be related to the number of online inferences alone, but a combination of considering children's reading profiles and other reader characteristics provided a better explanation. Therefore, we suggest that children in the reading profiles use the text approach that is most suitable for their cognitive abilities to remember the text.

Children in all profiles struggled to extract central information from expository texts. Possibly the role that inferences play in centrality for narrative and expository texts is different. We suggest that research in developing readers routinely needs to consider different text genres to better understand how online processes affect memory and ultimately learning from text.

Chapter 4

Individual differences in children's comprehension of temporal relations: Dissociable contributions of working memory capacity and working memory updating

This chapter has been published as:
Karlsson, J., Jolles, D., Koornneef, A., van den Broek, P., & van Leijenhorst, L. (2019). Individual differences in children's comprehension of temporal relations: Dissociable contributions of working memory capacity and working memory updating. *Journal of Experimental Child Psychology, 185,* 1-18.
doi:10.1016/j.jecp.2019.04.007

Individual Differences in Children's Comprehension of Temporal Relations: Dissociable Contributions of Working Memory Capacity and Working Memory Updating

Abstract

In two experiments, we examined 9- to 12-year-olds'comprehension and processing of two-clause sentences with a temporal connective (*before* or *after*) in sentence-medial or sentence-initial position. We obtained measures of individual differences in Working Memory capacity and Working Memory updating to test their contributions to comprehension. We measured the accuracy of children's responses to the questions "what happened first?" (Experiment 1; N = 74) or "what happened last?" (Experiment 2; N = 50) as well as their sentence reading times. Together, these experiments show continued development of comprehension of temporal relations in children in upper elementary school, and suggest that children's comprehension difficulties (i.e., more comprehension errors and longer reading times) were influenced by clause salience and recency effects rather than sentence chronology or the familiarity of the connective. Our findings are consistent with a memory resource-limited account and suggest that individual differences in WM updating and WM capacity make dissociable contributions to processing and comprehension of sentences with temporal order information.

Keywords: Clause Salience; Temporal Connectives; Recency; Reading Comprehension; Developmental Science

4.1 Introduction

Children in upper elementary school have to read and comprehend two-clause sentences describing a sequence of events on a daily basis. Sentences such as "Write down the answer to the question after you read this paragraph" or "Before you add or subtract you should solve the multiplication" are commonly used to instruct students at school. Cognitive theories of reading comprehension suggest that in order to understand these sentences, readers have to integrate the meaning of the individual words into a coherent mental representation (McNamara & Magliano, 2009; van Dijk & Kintsch, 1983). An essential aspect of creating a coherent representation is the processing of the temporal relations of the events represented in the text (Claus & Kelter, 2006; van den Broek, 1990; Zwaan, 1996). In sentences such as our example sentences above readers must use the temporal connectives *before* and *after* to establish the order in which the events represented in a sentence occurred and integrate this information in their emerging representation of the meaning of the sentence (Mann & Thompson, 1986; Van Silfhout, Evers-Vermeul, & Sanders, 2015). Comprehension of the correct temporal order between events in two-clause sentences is especially challenging because the clauses can occur in two orders. The sentences "You should solve the multiplication before you add or subtract" and "Before you add or subtract you should solve the multiplication" both describe the same order of events in the real world, however events are presented chronologically only in the first sentence.

Experimental studies have found that even though comprehension of spoken two-clause sentences starts developing in preschool and early elementary school (e.g. Blything, Davies, & Cain, 2015; Clark, 1971), comprehension of temporal relations in written two-clause sentences continues to improve throughout elementary school (Cain & Nash, 2011; Pyykkönen & Järvikivi, 2012). Different theoretical accounts have tried to explain these findings, but because the sample and methods used to examine these questions have varied between studies, our understanding of the circumstances that help or hinder comprehension of these type of sentences remains limited. Importantly, different factors might be involved during different stages of development. As prior research has focused mostly on preschoolers and early elementary school children, it is currently unclear what factors are most important during upper elementary school. Moreover, there are large individual differences in comprehension between children. A better understanding of the source of these differences could help educators identify and help those readers that need help the most.

Whereas in younger children semantic factors best explain comprehension difficulties for non-chronological sentences (de Ruiter, Theakston, Brandt & Lieven, 2018), recent studies suggest that a memory capacity-constrained theoretical account (e.g. Just & Carpenter, 1992) best explains performance differences in older children (Blything & Cain, 2016; de Ruiter, Theakston, Brandt & Lieven, 2018; Pyykkönen & Järvikivi, 2012). In the current study, we conduct two experiments and investigate the effect of textual factors (connective position, connective familiarity, and clause salience) as well as individual differences in Working Memory (WM) on the comprehension of temporal relations in 9-12-year-old children. We examine upper elementary school children because WM is still immature in this age range, while the demands on reading comprehension are relatively high. Additionally, prior studies suggest that comprehension of non-chronological sentences is more taxing, even for adults (Münte, Schiltz, & Kutas, 1998; Ye et al., 2012a, 2012b), and still immature in 12-year-olds (Pyykkönen & Järvikivi, 2012). More specifically we aim to further develop theoretical accounts of comprehension of temporal relations in upper elementary school readers by examining the influence of individual differences in both WM capacity and WM updating.

4.1.1 Working Memory and Reading Comprehension

The memory capacity-constrained framework (e.g. Just & Carpenter, 1992) explains difficulties that readers encounter in processing the temporal relations in non-chronological sentences in terms of the demand such sentences makes on their limited WM resources. Even though there are many different definitions of WM (for a review see Cowan, 2017), many converge on the notion that WM can be seen as a complex mental workspace in which one can keep, update, and manipulate information (e.g. Baddeley, 2003). In the context of language comprehension research, a definition of WM often entails both a limited-capacity store that holds information in a heightened state of availability, *and* the processes necessary to update the contents of this storage space, and is measured using complex WM span tasks (see e.g. Carretti, Borella, Cornoldi, & De Beni, 2009; Daneman & Merikle, 1996). WM develops throughout childhood and adolescence (e.g. Diamond, 2013; Gathercole, 1999; Huizinga, Dolan, & van der Molen, 2006). This development is thought to underlie the development of more complex cognitive abilities such as reading comprehension. Syntactically or semantically complex sentences such as two-clause sentences with temporal connectives require the reader to integrate different syntactical structures such as main and subordinate clauses in WM (e.g. Natsopoulos & Abadzi, 1986), and update their mental representation in response to linguistic cues. Comprehension of such sentences will be impeded if (1) the temporal sequence of events cannot be inferred from knowledge schemas that describe the typical order of the events (for example a sentence about a person who brushes her teeth before she goes to bed; e.g. French & Brown, 1977), and (2) if the reader does not have enough WM capacity and updating abilities to resolve the sequence of events based on textual cues. Hence, it could be argued that non-chronological sentences are more difficult compared to chronological sentences because participants have to use their WM to switch the order of events in their emerging mental representation. Crucially, theories within a memory capacity-constrained framework assume that individual differences in WM resources between readers can explain differences in comprehension of sentences containing temporal connectives.

Prior studies in 3-7-year-old children (Blything, Davies & Cain, 2015; Blything & Cain, 2016) and adults (Münte, Schiltz, & Kutas, 1998; Ye et al., 2012a, 2012b) have already shown that comprehension difficulties for non-chronological sentences were more pronounced in individuals with a more limited WM. For example, Blything, Davies and Cain (2015) and Blything and Cain (2016) found that individual differences in 3-7-year-olds comprehension of sentences with the temporal connectives *before* and *after* could be predicted by their performance on a simple WM span task. However, these studies were in younger children and used a listening comprehension, not a reading comprehension task. Furthermore, a number of studies in adults explained individual differences in brain activation in response to chronological and non-chronological sentences in terms of increased demands on WM for non-chronological sentences compared to chronological sentences (Münte, Schiltz, & Kutas, 1998; Ye et al., 2012a, 2012b). However, these studies did not include a behavioral measure of sentence comprehension.

Thus, there is some evidence that individual differences in WM influence comprehension of temporal relations in children and adults. However, the measures of WM that have been used in these prior studies do not allow for very specific conclusions, because they do not allow a disentangling of the roles of the development of WM capacity (the ability to store and process items in WM) and WM updating (the ability to continuously update and monitor the contents of WM). Moreover, to our knowledge, only two studies to date have examined comprehension of temporal relations in upper elementary school (Cain & Nash, 2011; Pyykkönen & Järvikivi, 2012). Cain and Nash (2011) compared 8-year-olds', 10-year-olds', and adults' comprehension of different types of connectives using online and offline tasks and found that 10-year-olds' ability to process and comprehend temporal connectives was still immature. Pyykkönen and Järvikivi (2012) found immature comprehension of two-clause sentences with temporal connectives in children up to 12 years of age. These studies with older children have not directly related comprehension to individual differences in WM. Our aim is to extend these prior studies by examining the relation between sentence comprehension and individual differences in WM in upper elementary school children. Moreover, for the first time, we will differentiate between WM capacity and WM updating abilities. Because prior studies in adults have shown that WM capacity and WM updating are not necessarily correlated (e.g. Jaeggi, Buschkuehl, Perrig, & Meier, 2010; but see Schmiedek, Hildebrandt, Lövdén, Lindenberger, & Wilhelm, 2009), and differentiation of WM has proven fruitful to better understand children's language comprehension in prior work (Finney, Montgomery, Gillam, & Evans, 2014). Using both a capacity- and an updating measure might elucidate how different aspects of WM influence comprehension of temporal relations in upper elementary school children.

4.1.2 Textual Characteristics Influencing Comprehension of Temporal Relations

Several textual factors have been found to influence comprehension of two-clause sentences with temporal connectives (e.g. Blything & Cain, 2016; de Ruiter, Theakston, Brandt & Lieven, 2018; Natsopoulos & Abadzi, 1986). In two-clause sentences, chronology results from the interaction of the temporal connective (*after* or *before*) and its position in the sentence (sentence-initial or sentence-medial). Results from previous studies appear contradicting regarding the role of the position of the connective in children. On the one hand, Blything and colleagues (2015) found that 3-7-year-old children who were asked to listen to two-clause sentences with temporal connectives and report the first event, made more errors on sentences with sentence-initial connectives than on sentences with sentence-medial connectives. The authors argued that a sentence-initial connective places a heavier load on comprehenders' WM capacity because the connective has to be kept in mind while attending to further information and can be used to order the events only at the end of the sentence. The additional demands that actively retaining this information places on WM may impoverish comprehension processes (Gibson, 2000; Gibson & Pearlmutter, 1998). On the other hand, Pyykkönen and Järvikivi (2012) found that 8-12-year-old children who were asked to read two-clause sentences with temporal connectives and report the first event, made more errors on sentences with sentence-medial connectives than on sentences with sentence-initial connectives. The authors suggest that sentence-medial connectives force readers to update the temporal order of events in their emerging mental representation in WM. Taken together, although both studies point to differences between conditions in the demands that are placed on children's limited WM resources, one study indicated that sentence-initial connectives are more difficult because they place a heavier load on readers' WM capacity (Blything et al., 2015), whereas the other suggested that sentence-medial connectives are more difficult because they demand readers to update their mental representations (Pyykkönen & Järvikivi, 2012). Critically, both studies did not directly test the relation between comprehension and WM with consideration of the influence of the position of the connectives.

Besides the position of the connectives, differences in familiarity of the connectives *before* and *after* itself could also be relevant. Even though semantic knowledge of temporal connectives seems to develop in early childhood (Blything & Cain, 2016; de Ruiter, Theakston, Brandt & Lieven, 2018; Evers-Vermeul & Sanders, 2009), differences in the use of the connectives *before* and *after* might continue to have an effect. For English speaking children, comprehension of *before* and comprehension of *after* have been suggested to develop differently due to a difference in linguistic complexity (e.g. Clark, 1971). In English the connective *after* is more ambiguous than the connective *before* because it is used in more ways than to indicate temporal relations (e.g. “Ben chased after a dog”, or “Ben chased the dog after he played with the cat”). Words that are more ambiguous are typically associated with processing costs (e.g. Eddington & Tokowicz, 2015). Although there is no such difference in Dutch and Dutch children commonly encounter both words, *before* (‘voordat’) is somewhat more frequent than *after* (‘nadat’) (Tellings, Hulsbosch, Vermeer, & van den Bosch, 2015). Therefore, we include this familiarity factor of the connectives in our analyses.

Finally, most prior research concerning two-clause sentences with temporal connectives has required participants to indicate which event happened first in a forced-choice task (e.g. Natsopoulos & Abadzi, 1986; Pyykkönen & Järvikivi, 2012; but see Blything & Cain, 2016). In doing so, the answer was always situated in the main clause when the connective *before* was used, and in the subordinate clause when the connective *after* was used. This is important because clause salience could influence comprehension. According to a syntactic account of the processing of complex sentences, certain syntactic configurations facilitate or hinder comprehension. Sentences are thought to be easier to comprehend if they follow a main-subordinate clause order (Diessel, 2005), and readers are more likely to attend to, encode, and remember the information in a main clause than in a subordinate clause (Cooreman & Sanford, 1996; Miltsakaki, 2003; Sanford, 2002). For example, adults are more likely to detect false statements when these false statements are presented in the main clause than when they are presented in the subordinate clause (Baker & Wagner, 1987). Clause effects on comprehension do not depend on extensive communicative experience and are found as early as in preschool. For example, to comprehend the temporal order between events in a listening task, preschool children treat information in the main clause as more important (Trosborg, 1982). This preference for main clauses seems especially strong when children cannot rely on their world knowledge to infer the temporal order (e.g. French & Brown, 1977). Moreover, by asking which event happened first the most recently read event corresponds to the correct answer for non-chronological sentences, not for chronological sentences. This could inflate accuracy scores for non-chronological sentences. Changing the comprehension question to "what happened last?" shifts the types of sentences for which comprehension of temporal relations is facilitated (see Table 4.1a and 4.1b), which allows us to investigate if the effects found in previous studies replicate with different task demands. Blything and Cain (2016) asked their participants to indicate what happened last and found lower accuracy overall compared to previous studies, which they attributed to these task differences. To better understand clause effects and recency effects the current study therefore includes two experiments: in the first experiment children are asked to indicate what happened first and in the second experiment they are asked to indicate what happened last.

In sum, research to date has suggested that children's difficulty in processing non-chronological sentences in upper elementary school can be best explained by a memory capacity-constrained account. However, very few studies to date examined children in upper elementary school, and the studies that did examine this age group did not include all textual factors that could influence online processing of temporal connectives, and did not directly measure WM. We aimed to close this gap in the literature by including measures of individual differences in WM capacity and WM updating in our study of processing and comprehension of two clause sentences with temporal connectives.

4.1.3 The current study

In the current study, we examined 9- to 12-year-old children's comprehension of temporal relations in sentences with two successive events and temporal connectives. We designed a computerized self-paced reading task based on the paradigm used by Pyykkönen and Järvikivi (2012) in which children read two-clause sentences containing the temporal connective *before* or *after,* in sentence initial or sentence medial position. This manipulation caused events to be presented either in chronological or non-chronological order (see Table 4.1). Following each sentence children were shown pictures of both events and were asked to indicate which of these events occurred first (Experiment 1), or last (Experiment 2) (see Figure 4.1). It should be noted that our study differed from the Pyykkönen and Järvikivi study in that the current computerized task did not allow for rereading: on each trial the sentence appeared on the first screen and was then replaced by a second screen with the comprehension question. As a consequence, our task likely puts greater demands on the participants' WM. The present task's demands on

WM are more similar to the listening task used by Blything and colleagues (2015) with preschool children.

Our first hypothesis, in line with previous studies was that 9- to 12-year-old children would perform above chance for all sentence types, reflecting a general understanding of the temporal connectives (Blything & Cain, 2016; Cain & Nash, 2011; Pyykkönen & Järvikivi, 2012). However, in line with previous studies we predicted that non-chronological sentences would be more difficult to process compared to chronological sentences because these sentences place greater demands on the reader's limited WM resources. These difficulties should be reflected in lower accuracy scores and longer reading times.

In addition, the current study is the first that aims to explore the differential roles of WM capacity and WM updating in comprehension of complex sentences with temporal connectives by directly examining individual differences in both types of WM ability. Therefore, we included a measure of WM capacity, a Sentence Span task (Swanson, Cochran, & Ewers, 1989), which measures participants' capacity to keep words in WM while performing a second task (i.e. answering a question), and a measure of WM updating, the Mental Counters task (Huizinga et al., 2006), which measures participants' ability to continuously update and monitor numerical information kept in WM. Following previous studies, our second and third hypotheses were that individual differences in WM updating would be related to comprehension of non-chronological sentences with a sentence-medial temporal connective (Pyykkönen & Järvikivi, 2012), because these sentences require the reader to update the order of information in their evolving mental representation of the sentences. In addition, we hypothesized that individual differences in WM capacity would be related to comprehension of sentences with a sentence-initial temporal connective (Blything et al., 2015) or of sentences in which the correct answer is situated in the subordinate clause, because both these factors place high demands on WM capacity as the reader has to actively retain the information in WM while also attending to other information.

Table 4.1.

Example sentences and comprehension questions for experiment 1 (a) and experiment 2 (b) and their relation to the factors that are hypothesized to facilitate comprehension of temporal relations.

	Main Clause	Connective position	Connective	Chronology
a. *Experiment 1. Comprehension question: "What happened first?"*				
Before Bart ate a cookie, **he drank milk**.	√	–	√	–
Bart ate a cookie before he drank milk.	√	√	√	√
After **Bart ate a cookie**, he drank milk.	–	–	–	√
Bart ate a cookie after **he drank milk**	–	√	–	–
b. *Experiment 2. Comprehension question: "What happened last?"*				
Before **Bart ate a cookie**, he drank milk.	–	–	√	–
Bart ate a cookie before **he drank milk**.	–	√	√	√
After Bart ate a cookie, **he drank milk**.	√	–	–	√
Bart ate a cookie after he drank milk	√	√	–	–

Note. For both experiments correct answers to the comprehension questions are shown in bold in the example sentence for each condition.
"√" indicates facilitation of comprehension and "-" indicate no facilitation by the correct response presented in the main clause, the position of the connective, the connective itself (*before* and *after*), or chronology (chronological vs. non-chronological sentences).

4.2 General Method

4.2.1 Materials

4.2.1.1 Reading Task

Comprehension of temporal relations was measured using a computerized reading task (E-prime version 2.0.8). On each trial, participants were asked to read a two-clause sentence describing two events and to determine which event occurred first (experiment 1) or which event occurred last (experiment 2). Eighty-four unique Dutch sentences were constructed, each representing two events that do not typically occur in a specific order. Thereby they preclude participants from relying on world knowledge so that they had to use the temporal connective to understand the sentences. Sentences contained one of three temporal connectives, *after* ('nadat'), *before* ('voordat'), or *while* ('terwijl'), which were presented in sentence-initial position or sentence-medial position. Our analyses focused on sentences with the connectives *after* and *before,* signaling a sequential order of the two events. The manipulation of these two connectives and their position in the sentence resulted in four sentence types (see Table 4.1a and 4.1b). Sentences with the connective *while,* indicating that events occurred simultaneously, were treated as filler trials. Participants were given 14 trials of each sentence type in a semi-randomized order, assuring the same type of sentences was not presented successively. The order of trials was counterbalanced between participants.

On each trial (see Figure 4.1), the sentence was followed by a screen on which the question "What happened first?" (Experiment 1) or "What happened last?" (Experiment 2) was presented at the top. Underneath the question, three pictures (from Microsoft Office ClipArt 2010) representing the response options were presented (see Figure 4.1). Text (in black), and pictures (in color) were presented on the screen against a light-grey background. The leftmost response option represented both events occurring simultaneously on each trial. The middle and right response options each represented one of the two events from the sentence. The position of the correct response in either the middle or the right picture was counterbalanced within conditions. The task was self-paced, after reading the sentence participants pressed the 'D' key on the keyboard with their left index finger to continue to the question. To choose the picture corresponding to the correct answer the participant pressed the 'A', 'S' or 'D' on the keyboard for response options 1, 2, or 3 with their left ring finger, middle finger or index finger, respectively. Each sentence and question remained on the screen for a maximum of 10 seconds. Between trials a fixation cross was presented for 500 ms.

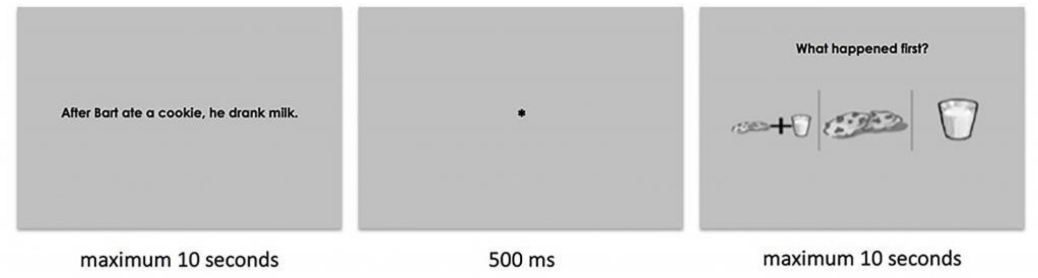

Figure 4.1. Schematic representation of a trial in the reading task. In this example two events (Bart ate a cookie, Bart drank milk) were shown with the sentence initial temporal connective after. Sentence presentation was self-paced with a maximum duration of 10 seconds. The sentence was followed by 500 ms fixation, which was followed by the comprehension question. Participants answered the comprehension question, which was presented at the top of the screen by selecting one of the three images that represent the choice options; "both events happened simultaneously" (which was always presented in the left panel), "Bart ate a cookie" (middle panel), and "Bart drank milk" (right panel). Within each condition, the position of the latter two response alternatives was counterbalanced between the middle and right panel.

A pilot study including 82 typically developing 9-year-old Dutch children (37 girls, M_{age} = 9 years, 4 months, SD_{age} = 4 months) was conducted to assess the age appropriateness of the reading task. Participants in the pilot study received a paper-and-pencil task in which they were asked to indicate the first event in each sentence by circling the correct answer. The overall mean accuracy score was fairly high: 80.17% (*SD* = 17.04), indicating the reading material was suitable for our test population. In addition, to test whether the pictures in the computerized task were easy to interpret a subset of the children who participated in experiment 1 (N = 49; 31 girls, M_{age} = 10 years, 7 months, SD_{age} = 1 year, 2 months) took part in a picture-word-matching test after completion of the main experiment. In this paper-and-pencil test, each of the 84 picture pairs used in the main experiment was shown to the participants, along with the corresponding words, and participants were asked to match each picture to the corresponding word. The overall mean accuracy of the picture-word-matching test was very high: 97.81% (*SD* = 4.92), indicating that the children understood the pictures.

4.2.1.2 Working-Memory Capacity

To assess participants' capacity to store items in WM while performing an interfering task we used a Dutch version of the Sentence Span Task by Swanson, Cochran, and Ewers (1989). The task contained four levels with two trials each. On each trial participants listened to a series of unrelated sentences and were asked to remember the last word of each sentence. After answering a comprehension question about one of the sentences, participants recalled the last word of each sentence. At the easiest level children listened to series of two sentences. At each level one sentence was added, increasing WM load, resulting in series of five sentences at the most difficult level. If the participant made an error on both trials within a level, the test was discontinued. Participants' accuracy scores, the sum of correctly remembered words for trials in which the question was answered correctly, were used in the analysis. This scoring method has gained a good internal consistency of .79 (Conway et al., 2005). The maximum score was 28.

4.2.1.3 Working-Memory Updating

To assess participants' ability to continuously update and monitor the stream of items held in WM we used the Mental Counters task. The Mental Counters task was developed by Huizinga and colleagues (2006) as part of a large executive function battery created to examine developmental changes in these functions. The task contained two sets of 20 trials each. On each trial participants were required to keep track of the score of counters presented visually as a black line on a white computer screen. Blocks appeared above or below the counters indicating if the score rose or fell with one point, respectively. The participant was required to indicate when the score for one of the counters reached above a certain number that was specified before the trial started. In the first set of trials participants had to track two counters, in the second set a third counter was added, increasing WM load. Participants' total accuracy score, the percentage of correct trials, was used in the analysis.

4.2.1.4 Raven Standardized Progressive Matrices

To obtain an estimate of general intellectual ability, participants completed a group-administered version of Raven's Standard Progressive Matrices (SPM: Raven, Raven, & Court, 1998). This measurement was included to test whether participants across the two experiments formed comparable groups of participants. Participants received a booklet with matrices, each one with a missing part, and were instructed to "solve as many puzzles" as possible within 30 minutes. The participants' task was to find the missing part out of six or eight options and write the number of the correct answer on an answering sheet. The items continuously increase in difficulty throughout the test. The estimated IQ scores were determined using international norms (Raven et al., 1998).

4.2.2 Procedure

Both experiments in the study were performed in two sessions. In the first session participants completed the Raven SPM in a classroom setting. In the month following the first session participants were tested individually at their schools. Each participant first performed the reading task, second the WM capacity task, and third the WM updating task. The reading task was explained using Powerpoint slides and six practice trials were completed in E-prime before the test trials began. The WM capacity task was explained verbally, and three practice trials were completed before the test trials began. The WM updating task was explained verbally with the use of visually presented examples, and three practice trials were completed before the test trials began. The reading task and the WM updating task were performed on 15-inch wide-screen Dell Latitude e6530 laptops running on windows 7 and an Intel core i5. The individual session took approximately 80 minutes to complete, including a break. The test battery included two further tasks (one in the first session and one in the individual session) which will, however, not be reported on in the current paper. At the end of the individual session, participants were thanked with a small gift. The study was carried out in accordance with the Ethical Declaration of Helsinki.

4.2.3 Analyses

To assess reading comprehension of temporal relations, sentence reading times and responses to the questions "What happened first?" (Experiment 1) and "What happened last?" (Experiment 2) were collected. The responses to the questions were binary scored as correct or incorrect. Generalized linear mixed-effects regression models were fitted for the accuracy data and linear mixed-effects regression models were fitted for the reading time data. The analyses were conducted with the statistical software R (version 3.5.1; R Core Team, 2018) using the package Lme4 (version 1.1-19; Bates, Mächler, Bolker, & Walker, 2015). Each model included the fixed factors Connective (*before* vs. *after*), Position (the connective appeared sentence-initial or sentence-medial), WM capacity, WM updating, and the full interactional terms of Connective by Position by WM capacity and Connective by Position by WM updating.[1] Participants and items were included as crossed random effects (Baayen, Davidson, & Bates, 2008). Sum coding was applied for the categorical independent variables (the connective *before* was coded

[1] R (Lme4) formula: Dependent variable ~ 1+Connective*Position*Sentence Span+ Connective*Position*Mental Counters+(1 | Participant)+ (1 | Item)

as -0.5 and the connective *after* was coded as 0.5; sentence-initial conditions were coded as -0.5 and sentence-medial conditions were coded as 0.5). The continuous predictors WM capacity and WM updating were centered to the mean. Fixed-effects estimates, standard errors, and the associated t-values (for the continuous dependent variable reading time) and z-values (for the categorical dependent variable accuracy) will be reported. To obtain p-values for the t-statistics, we follow the practice of Barr, Levy, Scheepers, and Tily (2013) and base those values on z-statistics as well.

4.3 Experiment 1

4.3.1 Method

4.3.1.1 Participants

Eighty-two typically developing children (50 girls) between the ages of 9 and 12 years (*M* = 10 years, 9 months, *SD* = 1 year, 1 month) were recruited from three primary schools located in middle-class neighborhoods in the Netherlands (Knol, 2012). Inclusion criteria were Dutch as mother tongue, and average to above-average scores on a standardized test on word reading developed by the Dutch Central Institute for Test Development (CITO; Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010). These scores were obtained from the schools. Children's scores on the Raven's SPM were in the average to above-average range (*M* = 113.13, *SD* = 11.06). Written parental consent was obtained for all children, and all participating children provided oral assent.

4.3.2 Results

Prior to all analyses, 8 participants' data were removed from the dataset because they were unable to complete one or more tasks (i.e., data was missing for the reading, WM capacity and/or WM updating tasks). For the remaining 74 participants (44 girls) we removed trials with reading times below 1000ms and trials with reading times above 10s. Similarly, trials in which the reaction time to the question was below 200ms or above 10s were removed from the analyses. As a result of this procedure 1.4% of the trials was removed. Table 4.2 and Figure 4.2 (top panel) report the mean accuracy scores (i.e., probability correct), reading times (in ms), and their standard errors (SE) as a function of the factors Connective and Position. On average, children obtained a score of 8.9 (range: 0-21) on the WM capacity task and a score of 81 (range: 58-100) on the WM updating task. The scores for the WM tasks did not show statistically significant correlations (r (72) = .13, p = 0.25).

Table 4.2

Mean accuracy scores (probability correct), reading times (in ms), and their standard errors (SE) in Experiments 1 and 2 as a function of Position and Connective.

		Experiment 1				Experiment 2			
Position	Connective	Accuracy	SE	Reading time	SE	Accuracy	SE	Reading time	SE
Initial	After	.74	.01	5563	65	.80	.02	5934	76
	Before	.93	.01	5100	63	.67	.02	5759	75
Medial	After	.85	.01	5289	63	.81	.02	5815	73
	Before	.93	.01	5201	59	.77	.02	5690	72

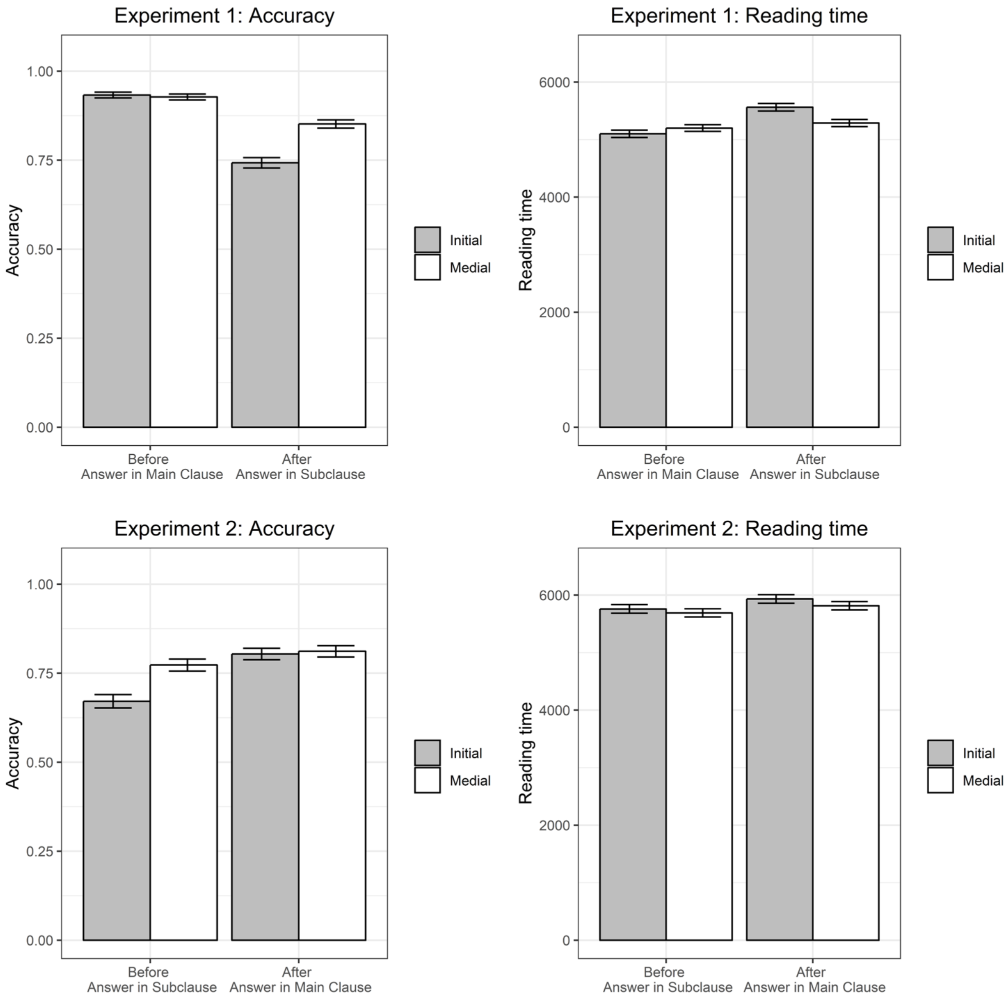

Figure 4.2. Mean accuracy scores (left) and reading times (right) in Experiments 1 (top panel) and 2 (bottom panel) as a function of Connective and Position. Error bars reflect standard errors (SE). Before-Medial and After-Initial sentences are chronological, Before-Initial and After-Medial sentences are non-chronological.

4.3.2.1 Accuracy

Table 4.3 reports the estimates (β-values), and the associated statistics for each fixed effect (excluding the Intercept) of the generalized linear mixed-effects regression model that was fitted to examine children's comprehension of temporal relations in Experiment 1. The analysis revealed a main effect of Connective, WM updating, and a Connective by Position interaction. For the fixed effect of Connective, the estimate carries a negative sign indicating that the accuracy scores were lower for sentences with the connective *after* (where the correct answer was situated in the subordinate clause) than for sentences with the connective *before* (where the correct answer was situated in the main clause). The main effect of WM updating carries a positive sign, indicating that children's comprehension accuracy scores increased as a function of how well they performed on the WM updating task. Figure 4.3 illustrates the Connective by Position interaction. Follow-up analyses (i.e., we fitted identical models, yet dummy-coded the categorical independent variables and adjusted the reference category to examine the relevant simple main effects) showed that there was no effect of Position for sentences with the connective *before* (when the correct answer was situated in the main clause) ($\beta = 0.16$, $SE = 0.32$, $z = 0.51$, $p = .61$). However, for sentences with the connective *after* (when the correct answer was situated in the subordinate clause), accuracy scores were lower for sentence-initial compared to sentence-medial conditions ($\beta = 0.81$, $SE = 0.29$, $z = 2.82$, $p < .01$). Furthermore, the simple main effects of Connective were significant for both sentence-initial and sentence-medial comparisons (*after-initial* vs. *before-initial*; $\beta = 1.91$, $SE = 0.31$, $z = 6.21$, $p <.001$; *after-medial* vs. *before-medial*: $\beta = 0.93$, $SE = 0.31$, $z = 3.05$, $p < 0.01$). To summarize, we did not find lower accuracy scores for non-chronological sentences, by contrast accuracy scores were particularly low for the condition in which the correct answer was situated in the subordinate clause and the connective appeared in the beginning of the sentence, even though in this condition events were presented chronologically.

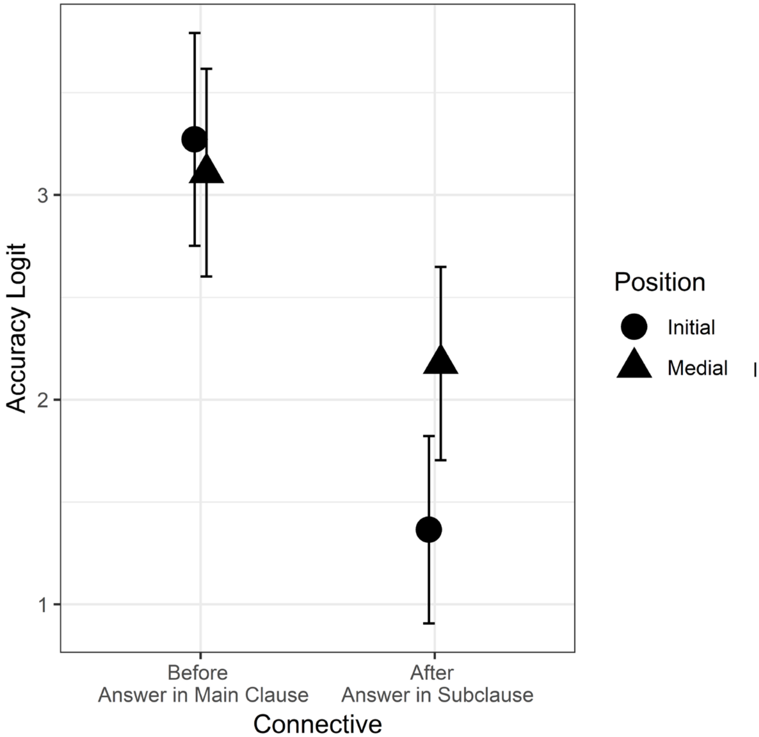

Figure 4.3. Fixed effects estimates (and their 95% confidence intervals) of the accuracy scores (logit scale) in Experiment 1 as a function of Connective and Position.

4.3.2.2 Reading Time

The linear mixed-effects regression analysis for the dependent variable reading time revealed main effects of Connective and WM capacity (Table 4.3). For the fixed effect Connective, the estimate carries a positive sign indicating that the reading times were longer for sentences with the connective *after* (where the correct answer was situated in the subordinate clause) than the reading times for sentences with the connective *before* (where the correct answer was situated in the main clause). The main effect of WM capacity carries a negative sign, indicating that children's reading times decreased as a function of how well they performed on the WM capacity task.

Table 4.3

Fixed effects estimates and associated statistics of the dependent variables in Experiment 1.

	Accuracy				Reading time			
Fixed Effects	*β*	*SE*	*z*	*p*	*β*	*SE*	*z*	*p*
Connective	-1.42	0.22	-6.53	<.01	282.51	134.38	2.10	.04
Position	0.32	0.22	1.50	.13	-76.93	134.38	-0.57	.57
WM capacity	0.04	0.03	1.14	.26	-52.70	24.07	-2.19	.03
WM updating	0.04	0.01	2.84	<.01	-12.97	10.63	-1.22	.22
Connective:Position	0.97	0.43	2.25	.02	-390.62	268.76	-1.45	.15
Connective:WM capacity	0.02	0.03	0.74	.46	-16.29	12.08	-1.35	.18
Position:WM capacity	0.02	0.03	0.60	.55	7.44	12.08	0.62	.54
Connective:WM updating	0.01	0.01	0.77	.44	1.57	5.43	0.29	.77
Position:WM updating	-0.01	0.01	-0.67	.50	-2.53	5.43	-0.47	.64
Connective:Position: WM capacity	-0.02	0.06	-0.43	.67	14.05	24.17	0.58	.56
Connective:Position: WM updating	0.04	0.02	1.63	.10	-2.72	10.85	-0.25	.80
Nr. of observations	3739				3739			
Nr. of participants	74				74			
Nr. of items	56				56			

4.3.3 Discussion

We tested how the temporal connectives *before* and *after* and the position of these temporal connectives affect 9-12-year-olds' comprehension of temporal relations between two events during reading. On average, comprehension scores were high, but there were performance differences between different sentence types. Our results differ from the results of previous studies because we did not find that processing and comprehension of non-chronological sentences was more challenging. By contrast, we found that children were most likely to give a wrong answer for sentences with the connective *after* – in which the correct answer was situated in the subordinate clause – especially when the connective was presented at the beginning of the sentence, even though this is a chronological sentence. Furthermore, reading times were longer for sentences where the correct answer was situated in the subordinate clause. These results indicate that clause salience could better explain task performance than sentence chronology. Interestingly, we found that on average, better WM updating was related to higher comprehension accuracy, whereas higher WM capacity was related to shorter reading times, suggesting that WM updating and WM capacity make separable contributions to sentence comprehension.

Experiment 1 did not allow us to rule out alternative explanations for these findings. First, facilitation of comprehension by the main clause was confounded with the familiarity of the connective. Therefore, our explanation that comprehension may have been compromised because the correct answer was situated in the less salient subordinate clause could be incorrect. An alternative explanation is that comprehension was compromised because the connective *after* is somewhat less common than the connective *before*. Second, our findings could be interpreted as a recency effect. Performance was best in the conditions in which the correct answer corresponded to the most recently read event (*before-initial* and *after-medial* sentences). To disentangle the effects of main clause salience and the familiarity of the connective, and to further examine recency effects on comprehension we conducted Experiment 2.

4.4 Experiment 2

In Experiment 2, we changed the comprehension question from "what happened first?" to "what happened last?". This modification changed the sentence types for which the correct response was facilitated by the main clause. Hence, in Experiment 2 the correct answer was located in the main clause for sentences with connective *after*, instead of the connective *before* as in Experiment 1. If comprehension is facilitated by the effect of salience of information in the main clause rather than the type of connective, one would expect that the sentence type for which most comprehension errors are made should shift from sentences with the sentence-initial connective *after* in Experiment 1 (see Table 4.1a), to sentences with the sentence-initial connective *before* in Experiment 2 (see Table 4.1b). If comprehension is facilitated by a recency effect, one would expect that performance would be better for sentences in which the correct answer corresponds to the most recently read event (*before-medial* and *after-initial* sentences).

4.4.1 Method

4.4.1.1 Participants

Fifty-three typically developing children (33 girls) between the ages of 9 and 12 years (*M* = 11 years, *SD* = 1 year, 1 month) were recruited from two primary schools located in middle-class neighborhoods in the Netherlands (Knol, 2012). Inclusion criteria were Dutch as mother tongue, and an average to above-average score on the standardized measure of word reading (CITO; Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010) provided by the school. Children's scores on the Raven's SPM were in the average to above-average range (*M* = 111.13, *SD* = 12.97) and did not differ significantly from the scores of children participating in Experiment 1 (t (130) = 0.94, p = .35). Written parental consent was obtained for all children, and all participating children provided oral assent.

CHAPTER 4

4.4.2 Results

Prior to all analyses, 2 participants' data were removed from the dataset because we were unable to collect the data for the WM tasks. Data from one additional participant was excluded because of failure to comply with the task instructions. For the remaining 50 participants (31 girls) we removed trials with reading times below 1000ms and trials with reading times above 10s. In addition, trials in which the reaction time to the question was below 200ms or above 10s were removed (13% of the trials was removed in total). Mean accuracy scores, reading times, and their standard errors are reported in Table 4.2 and Figure 4.2 as a function of the factors Position and Connective. On average, the children obtained a score of 8.0 (range: 0-21) on the WM capacity task and a score of 86 (range: 65-100) on the WM updating task. These scores did not correlate (r (48) = .17, p = 0.23).

4.4.2.1 Accuracy

The analysis revealed a main effect of Connective, a main effect of WM updating, a Connective by WM updating interaction, and a Connective by Position by WM updating interaction (see Table 4.4). The positive β-value for the fixed effect Connective indicates that the accuracy scores were higher for sentences with the connective *after* (where the correct answer was situated in the main clause) than for sentences with the connective *before* (where the correct answer was situated in the subordinate clause). This replicates the 'main-clause advantage' observed in Experiment 1. The positive β-value for the fixed effect WM updating shows that children's reading accuracy scores increased as a function of how well they performed on the WM updating task. Figure 4.4 depicts the significant two- and three-way interactions. The graph on the left shows that children with a higher score on the WM updating task are more sensitive to the main-clause advantage than children with a lower score on the WM updating task. However, as can be observed in the dual graph on the right (Figure 4.4), this effect interacts with the position of the connective as well. Two series of follow-up analyses were conducted to further interpret this three-way interaction.

First, we examined Position by WM updating interactions separately for the connectives *after* and *before* (i.e., we fitted models containing all the fixed effects of the original model, yet dummy coding was applied to obtain the relevant simple main effects). These analyses revealed that WM updating significantly interacts with the position of the connective *after* ($\beta = 0.052$, $SE = 0.019$, $z = 2.70$, $p < .01$), but not with the position of the connective *before* ($\beta = 0.016$, $SE = 0.016$, $z = 0.98$, $p = .33$). Moreover, only the *after-medial* condition showed a significant (positive) effect of WM updating, indicating that WM updating positively affects comprehension of a non-chronological sentence with a medial connective ($\beta = 0.083$, $SE = 0.020$, $z = 4.13$, $p < .001$; for all other conditions $z < 1.85$).

Second, to examine more directly how the factors Connective and Position affected children with lower and higher scores on the WM updating task, we adjusted the baseline value of this continuous predictor to the 1st (low) and 9th (high) decile (see vertical dotted lines in Figure 4.4). The analyses for children with a higher score on the WM updating task revealed a main effect of Connective only, indicating that children with a higher WM updating score performed better for sentences with the connective *after* (when the answer was situated in the main clause) ($\beta = 0.94$, $SE = 0.25$, $z = 3.75$, $p < 0.01$). The analyses revealed a different picture for children with a lower score on the WM updating measure. For these children there were no main effects of Connective and Position, yet the Connective by Position interaction was significant ($\beta = 1.32$, $SE = 0.50$, $z = 2.62$, $p < .01$). More specifically, for sentences with the connective *before* (when the correct answer was situated in the subordinate clause), performance was better for sentences with a sentence-medial connective than for sentences with a sentence-initial connective ($\beta = 0.76$, $SE = 0.35$, $z = 2.17$, $p = .03$). Numerically, the opposite pattern was present for sentences with the connective *after* (when the correct answer was part of the main clause), but this main effect was not significant ($\beta = 0.56$, $SE = 0.36$, $z = 1.54$, $p = .12$).

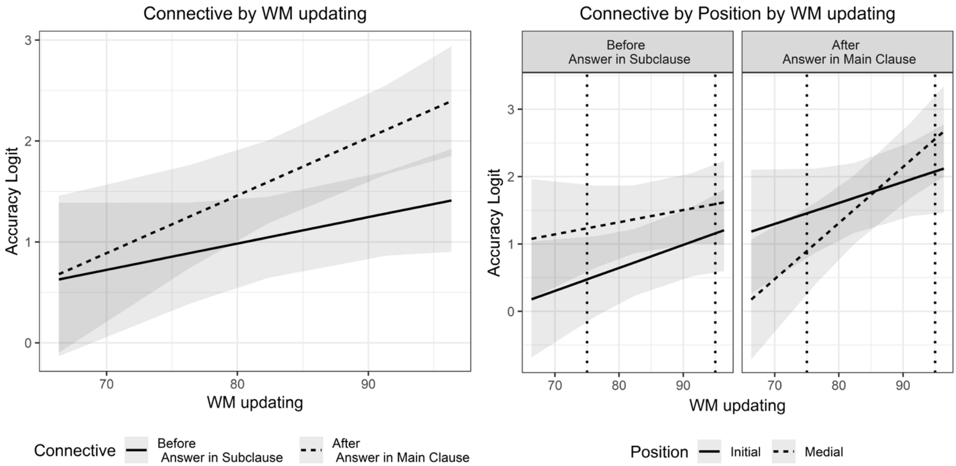

Figure 4.4. Fixed effects estimates (and their 95% confidence intervals) of the accuracy scores (logit scale) in Experiment 2 as a function of Connective, Position, and WM updating. The single graph on the left depicts the two-way interaction of Connective by WM updating and the dual graph on the right depicts the three-way interaction of Connective by Position by WM updating. The vertical dotted lines in the dual graph reflect WM updating scores at the 1st and 9th decile. Before-Medial and After-Initial sentences are chronological, Before-Initial and After-Medial sentences are non-chronological.

CHAPTER 4

4.4.2.2 Reading Time

The analysis revealed a main effect of WM capacity, a Position by WM capacity interaction, and a Position by WM updating interaction (see Table 4.4). The main effect of WM capacity carries a negative sign, indicating that children's reading times decreased as a function of how well they performed on the WM capacity task. Figure 4.5 depicts the two-way interactions and reveals a clear picture. The WM capacity score is a better predictor of the reading times (i.e., a higher WM capacity score is associated with shorter reading times) when the connectives appear sentence-initially than when they appear sentence-medially (see graph on the left in Figure 4.5). The opposite holds for the WM updating task. The score for this task is a better predictor of the reading times when connectives occur sentence-medially, as opposed to sentence-initially (see graph on the right in Figure 4.5).

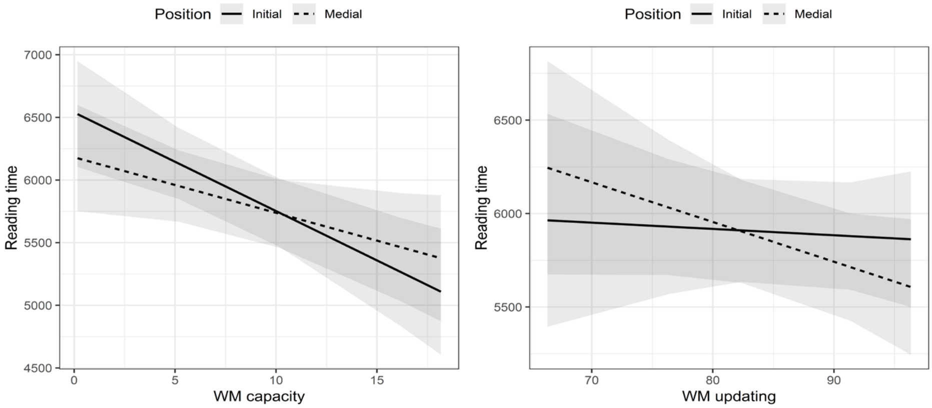

Figure 4.5. Fixed effects estimates (and their 95% confidence intervals) of the reading times in Experiment 2 as a function of WM capacity, WM updating, and Position. The graph on the left depicts the two-way interaction of WM capacity by Position and the graph on the right depicts the two-way interaction of WM updating by Position.

Table 4.4

Fixed effects and associated statistics of the dependent variables in Experiment 2.

	Accuracy				Reading time			
Fixed Effects	*β*	*SE*	*z*	*p*	*β*	*SE*	*z*	*p*
Connective	0.67	0.22	3.09	<.01	151.59	124.85	1.21	.22
Position	0.30	0.22	1.40	.16	-77.03	124.83	-0.62	.54
WM capacity	0.03	0.03	1.13	.26	-61.62	20.73	-2.97	<.01
WM updating	0.04	0.02	2.58	.01	-12.26	12.37	-0.99	.32
Connective:Position	-0.54	0.43	-1.25	.21	-71.17	249.70	-0.29	.78
Connective:WM capacity	-0.02	0.02	-1.09	.27	-11.77	13.36	-0.88	.38
Position:WM capacity	-0.03	0.02	-1.51	.13	34.34	13.36	2.57	.01
Connective:WM updating	0.03	0.01	2.45	.01	13.45	8.10	1.66	.10
Position:WM updating	0.02	0.01	1.42	.16	-18.01	8.11	-2.22	.03
Connective:Position: WM capacity	0.01	0.04	0.15	.88	-13.69	26.72	-0.51	.61
Connective:Position: WM updating	0.07	0.03	2.69	.01	-29.98	16.21	-1.85	.06
Nr. of observations	2443				2443			
Nr. of participants	50				50			
Nr. of items	56				56			

4.4.3 Discussion

The results of Experiment 2 replicated the results from Experiment 1. As in experiment 1, comprehension scores were high on average, but there were performance differences between different sentence types. Accuracy was a little lower in the second experiment. This is consistent with previous reports of slightly worse comprehension in response to the question "what happened last" reported by Blything and Cain (2016) in younger children. This increase in task difficulty may have caused individual differences between children to be more clear, and Experiment 2 revealed a more nuanced picture of how WM updating and WM capacity are differently related to processing and comprehension of two-clause sentences with temporal connectives.

We again did not find a clear effect of sentence chronology, in contrast, we replicated our finding that 9-12-year-old children were least likely to give the correct answer when it was presented in the subordinate clause. Hence, our prediction for Experiment 2 that the most difficult sentence type would shift and involve the connective *before* rather than *after* was correct. We did not find support for the explanation that a difference in familiarity between (the Dutch equivalents of) the words *before* and *after* affects comprehension in upper elementary school children (Blything & Cain, 2016; de Ruiter, Theakston, Brandt & Lieven, 2018; Evers-Vermeul & Sanders, 2009). Instead, our results support the interpretation that task-relevant information presented in the main clause facilitates comprehension across experiments. This finding is consistent with the interpretation that in older children comprehension of sentences containing temporal connectives cannot be explained by semantic differences between these sentences (Clark, 1971), and fits with a memory-capacity constrained framework (Just & Carpenter, 1992).

Interestingly, interactions with WM updating suggest comprehension differences between children as a function of their WM updating abilities. Taken together, the follow-up analyses for the interaction of the position of the connective with WM updating suggest that**:** First, WM updating is particularly relevant when the connective is in a sentence-medial position *and* the correct answer is part of the main clause. Second, children with a higher WM updating score show a main-clause advantage (i.e., they perform better when the correct answer is situated in the main clause). Third, children with a lower WM updating score do not display a main-clause advantage. Instead, they seem to perform better on chronological sentences, when the sentence structure follows the temporal sequence of events in the real world (which is the case in the *before-medial* and *after-initial* conditions, but not in the *before-initial* and *after-medial conditions*). At first glance this seems like an effect of sentence chronology, however it could also be interpreted as a recency effect. In sentences with a *before-medial* and *after-initial* connective the most recently read information is also the correct answer to the question "What happened last?".

Individual differences in WM ability were most apparent in our reading times results. While accuracy scores varied as a function of WM updating, reading times varied as a function of WM updating and WM capacity. Results showed both an interaction between the position of the connective and WM capacity and between the position of the connective and WM updating. WM capacity predicts reading times of sentences with sentence-initial connectives, supporting the hypothesis that these sentences place high demands on WM capacity as the reader has to actively retain the information in WM while also attending to other information. WM updating, on the other hand, predicts reading times of sentences with sentence-medial connectives, supporting the hypothesis that these require the reader to update the order of information in the evolving mental model of the sentence.

4.5 General Discussion

The aim of the two experiments in this study was to examine upper elementary school children's comprehension of the temporal relations between two events while reading two-clause sentences. As expected, sentence comprehension was relatively good for these 9-12-year-old children in both experiments. Even though our main results did not replicate previous findings that sentence chronology is an important factor influencing comprehension, our results are consistent with a memory capacity-constrained account of processing of these types of sentences in children (Blything & Cain, 2016; Blything, Davies & Cain, 2015) and adults (Münte, Schiltz, & Kutas, 1998; Ye et al., 2012a, 2012b). More specifically, we found that text factors that facilitate processing by reducing the demands on WM resources resulted in better comprehension. Interestingly, WM updating and WM capacity contributed differently to task performance. The subtle differences in accuracy between the two experiments, together with the finding in Experiment 2 that children with relatively poor WM updating abilities were most sensitive to task effects underline the importance of examining the interaction between textual factors and reader characteristics that influence comprehension in older children. In addition, our findings suggest that in upper elementary school comprehension of temporal relations in complex sentences is still not fully proficient, which fits with the results reported by Pyykkönen and Järvikivi (2012).

Children's comprehension of temporal connectives in our experiments was best explained by the effect of clause salience and the recency of the information that was the correct answer to the question in each task. Information in the subordinate clause proved to be especially difficult for children to comprehend. This finding is in line with previous studies in both adults (e.g. Baker & Wagner, 1987) and children (French & Brown, 1977; Trosborg, 1982), and could suggest that this syntactic factor contributes to comprehension in older children. However, it could also be interpreted in light of the demands these different types of sentences make on WM resources. In the current study the subordinate clause always starts with a temporal connective. As argued above, processing temporal relations and temporal connectives is an essential aspect of creating a coherent representation of a text (e.g. Claus & Kelter, 2006; Mann & Thompson, 1986; van den Broek, 1990; Van Silfhout et al., 2015; Zwaan, 1996). However, a connective may also be interpreted by a reader as a signal that the main idea is elsewhere and cause them to allocate more attention, and commit to deeper processing of the main clause rather than the subordinate clause (e.g. Sanford, 2002). In this context our finding that children with relatively good and poor WM updating abilities seem to approach these sentences differently is intriguing and adds new information to the literature on sentence comprehension in childhood. Children with good WM-updating abilities are sensitive to facilitating effects of the main clause and perform better when the correct answer coincides with the main clause, whereas comprehension in children with poor WM-updating abilities resembles the finding in younger children. For them chronological sentences seem easier to comprehend (e.g. Blything & Cain, 2016). However, the alternative explanation that this chronology effect is in reality a recency effect could also be true. Even though chronology and recency are confounded in Experiment 2, they are not in Experiment 1. Taken together, the results from both experiments suggest that recency effects might be a better explanation for the processing and comprehension of temporal connectives, at least in older children. Future work should explore these alternative explanations, as well as examine possible task-related effects.

Reading times largely mirrored the accuracy results, consistent with the interpretation that better comprehension is a consequence of reduced demands on WM resources. Thus, our reading times results support a memory capacity-constrained account as well. In both experiments better WM capacity was related to faster processing of sentences. In Experiment 2 – where slightly lower comprehension scores might have allowed us to better capture individual differences – WM capacity and WM updating both interacted with the position of the connective. On the one hand a bigger WM capacity facilitated processing of sentences with sentence-initial connectives, while on the other hand better WM updating facilitated processing of sentences with sentence-medial connectives. These findings reconcile previous seemingly contradictory findings that sentence-initial connective place high demands on WM resources because they require the reader to keep this information in mind while processing incoming information (Blything et al., 2015), and that sentence-medial connectives place high demands on WM resources because these require the reader to update the order of information in the evolving mental model of the sentence in (Pyykkönen & Järvikivi, 2012). These findings demonstrate the importance of using several measures of WM in developmental studies to unravel the relation between reading comprehension processes and individual differences.

Our study focused on 9-12-year-old readers in upper elementary school. In doing so we extended the literature by showing that in these children comprehension of connectives is not fully proficient which suggests that it continues to develop throughout adolescence. However, it should be noted that to further develop theoretical accounts of developmental change in readers comprehension of temporal relations, future studies should examine the influence of individual differences in both WM capacity and WM updating across a broader age range. In addition, in the context of this experiment readers encountered complex sentences with temporal connectives in isolation. It could be argued that demands on WM are different when these sentences are encountered in a text. Furthermore, we did not manipulate WM load directly in our experiment. An important next step would be to examine the effects of individual differences in response to changes in the demands made on WM capacity and WM updating abilities by the reading task. Nevertheless, our findings are an important initial step.

In conclusion, our findings show continued development of comprehension of temporal connectives in children in upper elementary school. Our findings are consistent with a memory resource-limited account and suggest that individual differences in WM updating and WM capacity make dissociable contributions to processing and comprehension of sentences with temporal order information. Therefore, our findings contribute to the further refinement of models of the development of the ability to comprehend temporal connectives during reading. It could be argued that the sentences that children encounter in everyday life are more complex than the sentences that we used in the present study. However, our findings suggest that even under these relatively simple circumstances comprehension of sentences with temporal connectives continues to show improvements until at least the age of 12.

Chapter 5

The Role of Working Memory in Inference Generation during Reading and Listening Comprehension

Ganushchak, L, Y., Karlsson, J., & van den Broek P. (in preparation). The role of working memory in inference generation during reading and listening comprehension.

The Role of Working Memory in Inference Generation during Reading and Listening Comprehension

Abstract

We investigated the role of working memory (WM) during reading and listening comprehension. The same participants read (eye-tracking) and listened (EEG) to short stories. A target word had either high predictability or low predictability, depending on the prior discourse context. Readers and listeners showed a typical predictability effect, with longer total gaze durations, more regressions, and more pronounced N400 amplitudes for low predictable targets words than high predictable target words. This effect was modulated by WM during listening but not during reading. In the listening task, low-WM participants but not high-WM participants showed an effect of predictability. We conclude that working memory may be more taxed during listening than during reading comprehension. Making predictions from local context alone and inhibition difficulties are possible mechanisms discussed as underlying the predictability effect in low-WM participants during listening comprehension.

Keywords: reading comprehension, listening comprehension, working memory, N400, predictive inferencing.

5.1 Introduction

The ability to comprehend written and spoken language is essential for an individual's ability to function in a modern information society. It is important to understand not only the meaning of isolated bits of information, but also the relations between them. During reading and listening, the successful comprehender connects encountered events, persons, facts, and objects so that the text appears to be coherent (e.g., van den Broek, Risden, Fletcher, & Thurlow, 1996). Working memory (WM), a system responsible for active maintenance of information during ongoing language processing (Conway et al., 2005), has been implicated as a major factor in the success in creating coherence and as a source of individual differences (e.g., Calvo, 2001; Daneman & Carpenter, 1980; Otten & van Berkum, 2009; Whitney, Ritchie, & Clark, 1991). For instance, Just and Carpenter (1992) suggested that an individual's working memory capacity provides a central constraint on general language comprehension. However, it is possible that written and spoken information make different demands on WM. For instance, listening requires that a comprehender holds each preceding word in working memory, whereas reading provides memory support in the form of written words that the comprehender may re-read. Consequently, the role of working memory may differ for these two comprehension modalities. Here we investigate how individual differences in working memory affect reading and listening comprehension.

5.1.1 Working Memory

Working memory is a multi-component system responsible for active maintenance of information in the face of ongoing processing and/or distractions (Conway et al., 2005). Arguably, the most influential model of WM is the one proposed by Baddeley (for a review see Baddeley, 2012). According to this model, an attentional control mechanism (i.e., the central executive) oversees three auxiliary domain-specific subsystems responsible for the temporary maintenance of verbal (the phonological loop), visual (the visuospatial sketchpad), and long-term memory information (Baddeley, 2012). Verbal information is maintained for short periods by the phonological loop whose contents decay over time unless it is maintained by a process of subvocal articulation (Baddeley, 2007) or via the focusing of attention onto different components of long-term memory (LTM; Cowan, 1995; Oberauer, 2003). Working-memory processes involved in maintaining information relevant to the task are inhibition, switching, and updating (although the exact relation amongst them are debated; Diamond, 2013; Ecker, Lewandowsky, Oberauer, Chee, 2010). The extent to which information maintenance depends on domain-specific skills versus domain-general executive attention varies as a function of individual ability, task context, and interaction between ability and context (Conway et al., 2005).

5.1.2 Inferences during Language Comprehension

Central to successful comprehension is the construction of a coherent mental representation of the provided information, a so-called mental model or a situation model (e.g., Graesser et al., 1997; Kintsch, 1988; Rapp, van den Broek, MacMaster, Kendeou, & Espin, 2007; Van den Broek, 2010). The processes of inference generation (i.e., extracting implicit information from the text) are important for the construction of such a model. Readers and listeners use discourse information and their world knowledge to make predictive inferences about upcoming events (e.g., Cook, Limber, & O'Brien, 2001; McKoon & Ratcliff, 1986; Weingartner, Guzmán, Levine, & Klin, 2003), but also to make predictions on the word level (e.g., activate aspects of word meaning, grammar, and form) prior to the onset of a predicted word (e.g., Delong, Urbach, & Kutas, 2005; Otten, Nieuwland, & van Berkum, 2007). It is believed that prediction plays an important role in effective use of language (e.g., Altmann & Mirković, 2009), and that predictions help readers and listeners to anticipate the forthcoming information and therefore facilitate subsequent processing (e.g., Calvo, 2001; Federmeier & Kutas, 1999a, 1999b; Fincher-Kiefer, 1995; Keefe & McDaniel, 1993; Klin, Murray, Levine, & Guzmán, 1999; Schwanenflugel & LaCount, 1988; Weingartner et al., 2003; Whitney, Ritchie, & Crane, 1992). Making predictions, however, depends on the text being constraining enough (Cook et al.., 2001; Klin, Guzmán, & Levine, 1999; van den Broek, 1990) as well as the comprehenders' characteristics, such as WM capacity (e.g., Linderholm, 2002; Virtue, van den Broek, & Linderholm, 2006).

5.1.3 Comprehension and Working Memory

According to the causal inference process model (van den Broek, Fletcher, & Risden, 1993) and the capacity constrained comprehension theory (Just & Carpenter, 1992), (predictive) inferences are made when new information initiates a spreading of activation to associated concepts from background knowledge stored in long-term memory (LTM). Associated concepts are transferred from LTM into WM provided that there is enough activation, or cognitive resources, for further processing. Predictive inferences are made when explicit and associated concepts co-occur in WM and connections are made between them. Thus, enough cognitive resources are needed to keep the meaning of previous words active as the discourse proceeds and, at the same time, combine these with associated concepts to generate the predictive inferences. Accordingly, the more working memory capacity that readers have available, the more likely it is that they will be able to expand on language processing and draw predictive inferences (Calvo, 2004; Just & Carpenter, 1992; Singer & Ritchot, 1996). For example, when reading short paragraphs requiring inference making for successful comprehension, high-WM participants were said to be more successful in generating predictive inferences than low-WM participants, as reflected by their N400 amplitudes (St George et al., 1997). The ERP component N400 is typically related to lexical and semantic processing (for a review, see Kutas & Federmeier, 2011). However, other research finds the N400 effect in both high- and low- WM participants (Otten & van Berkum, 2009), suggesting that both groups successfully make predictions. Instead, there may be qualitative differences in inference generation between comprehenders with different WM capacity. For instance, high- and low-WM participants appear to use different sources to generate an inference. High-WM participants rely both on discourse context and background knowledge to generate inferences from a larger pool of activated concepts, whereas low-WM participants tend to rely more on local discourse context (Boudewyn, Long, & Swab, 2013). Furthermore, low-WM but not high-WM participants may have difficulties in processing faulty predictions, shown by larger negativity at later processing (Otten & van Berkum, 2009). Low-WM participants have more problems with inhibiting unwanted information, compared to high-WM participants (e.g., Engle, 2002; Gernsbacher & Robertson, 1999), possibly explaining difficulties in processing faulty predictions.

5.1.4 Readings versus Listening Comprehension

Although it is clear that WM affects language comprehension, it is possible that this effect differs between spoken and written language. Previous studies have shown that listening and reading comprehension are two closely related skills (Booth, Perfetti, & MacWhinney, 1999; Booth, Perfetti, MacWhinney, & Hunt, 2000; Just & Carpenter, 1987). It is believed that higher-order cognitive processes (e.g., inference making) of text comprehension are modality independent and therefore the same in reading and listening comprehension (e.g., Booth et al., 2002). However, neuroimaging studies suggest that comprehension in these two modalities involves both similar and different underlying cognitive processes even for higher-level comprehension processes. Besides overlapping brain regions active during both reading and listening, there are also distinct brain regions that are only active during listening or reading comprehension (e.g., Buchweitz, Mason, Tomitch, & Just, 2009; Michael, Keller, Carpenter, & Just, 2001). Typically, listening comprehension results in more overall activation across the whole brain, whereas brain activation associated with reading comprehension is more lateralized in the left hemisphere (Buchweitz et al., 2009; Constable et al., 2004; Jobard e al., 2007; Michael et al., 2001). One of the explanations of this pattern is that listening comprehension is sequential in nature and therefore more demanding than reading comprehension, where the reader can re-read words. Typically, more demanding tasks result in more widespread brain activation than easier tasks, for example due to recruitment of areas important for executive control or information maintenance (e.g., Buchweitz et al., 2009; Constable et al., 2004; Michael et al., 2001). If listening poses higher cognitive demands on the comprehender than reading, then it is possible that the role of WM is more important in listening than in reading comprehension. The primary focus of the current study is to determine whether individual differences in WM affect prediction equally in the two comprehension modalities, reading and listening.

5.1.5 Current Study

We investigated the effects of individual differences in working memory on reading and listening comprehension using a task in which the ease of making predictive inferences was manipulated. Ease of making predictions was manipulated by varying expectancy of the target word. The same participants read and listened to short two-sentence stories. The first sentence provided the comprehender with a setting and context. In the second sentence we manipulated a target word to have either high or low predictability, depending on the discourse context provided by the first sentence. In example (1), the discourse context is set-up in such a way that *aquarium* rather than *bowl* is a more predictable word. In example (2) the reverse is true: *bowl* is more predictable than *aquarium*.

(1) "Peter vindt tropische vissen heel erg mooi. Thuis heeft hij veel verschillende soorten in een aquarium/kom zwemmen".
(Peter thinks that tropical fish are very beautiful. At home, he has a lot of different types swimming in an aquarium/bowl.)

(2) "Het jongetje was ontzettend blij met zijn nieuwe goudvis. Thuisgekomen deed hij de vis meteen in een kom/aquarium met schoon water."
(The boy was very happy with his new goldfish. At home he immediately put the fish in a bowl/ an aquarium with fresh water.)

To study cognitive processes during listening, we measured Event-Related Potentials (ERP) by means of electroencephalography (EEG). The ERP component of interest here is the N400. The N400 reflects the degree to which retrieval of semantic memory associated with a word is facilitated by the discourse context (for a review, see Kutas & Federmeier, 2011). Recently, it has been argued that the N400 also reflects the ability to use background knowledge to predict upcoming words (e.g., Nieuwland, 2015). Overall, a reduced N400 amplitude is found on words that are semantically plausible, related, or predictable given the preceding language context compared to words that are semantically implausible, unrelated, or unpredictable (e.g., Federmeier & Kutas, 1999a, 1999b; Kutas & Hillyard, 1980; van Berkum, Hagoort, & Brown, 1999; van Petten, 1993; for a review, see Swaab, Ledoux, Camblin, & Boudewyn, 2012). Typically, the N400 effect observed in sentence processing shows that both readers and listeners immediately relate the incoming words to a semantic representation of the preceding language input (e.g., van Berkum et al., 1999).

Using ERPs during a listening task is suitable as the stimuli can be presented as natural language. However, when using ERPs in reading research, sentences are typically presented word-by-word, and thus pose a potentially higher cognitive load because participants need to remember each word and cannot look back. A word-by-word presentation may therefore overestimate the role of WM during reading. This does not make ERPs suitable to address this paper's research questions about reading processes because it would require changing those natural aspects of reading that make reading processes possibly different from listening processes. Thus, we chose two different methods to study discourse processing (and the role of WM) in the two modalities, each of which is eminently suited to record on-line processing in one of the modalities.

To study cognitive processes during reading we used an eye-tracking methodology, as it allows participants to read as naturally as possible. I.e. they could read at their own pace and preview as well as re-read parts of the text. For the eye-tracking data, a distinction between early and late processing was made. The index of early processing was (a) first gaze duration, i.e., the sum of all first-pass fixations durations on a target word prior to moving to another word. Indices of late processing (e.g., re-analysis) were (b) probability of regressions, i.e., backward eye movements from a region back to the target word, and (c) total gaze duration, i.e., the sum of all fixations on a target word. Prior research has shown that first gaze duration is sensitive to early comprehension processes, such as word recognition (for an overview see Clifton, Staub, & Rayner, 2007). Total gaze durations reflect late processing, such as re-analysis and discourse integration (e.g., Frisson & Pickering, 1999; Rayner, 1998; Sturt, 2007). For both first gaze and total gaze measures, longer durations are interpreted as an indication for more effortful integration processes (e.g., Rayner & Sereno, 1994). Typically, readers tend to look longer at and have more regressions (re-reading of earlier parts of the text) to words that have low semantic plausibility or low predictability than words that have high semantic plausibility or high predictability considering the context (e.g., Pickering & Traxler, 1998; Rayner, Warren, Juhasz, & Liversedge, 2004; Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007; Warren, McConnell, & Rayner, 2008). This effect is believed to reflect readers' difficulty to generate semantically coherent interpretations of the sentences.

The current study employs the Mental Counters task to evaluate working memory (Huizinga, Dolan, & Van der Molen, 2006; Larson, Merritt, & Williams, 1988). In the studies reviewed above working memory was measured by the Reading Span task (Just & Carpenter, 1992) or by the Sentence Span task (Swanson, 1992). Both these tasks are complex span tasks that involve a tradeoff between maintenance and processing. Typically, participants are asked to read or listen to sets of unrelated sentences and to remember the final word of each sentence. Prior to recalling the final word of each sentence, participants are asked a question about one of the sentences. Thus, performance on such tasks is dependent on participants' comprehension ability. As our goal is to examine the relation between WM and listening and reading comprehension, a WM task independent of either modality is necessary. In the Mental Counters task participants are required to keep track of and update the score of visual counters in working memory, hence, performance on this task is independent from language comprehension ability.

We conducted the experiments to investigate (a) whether the role of WM differs during reading and listening comprehension, and (b) whether the inference generation process is comparable between the two modalities. If listening comprehension poses higher cognitive demands than reading comprehension, then a larger WM should be of more importance during listening than reading comprehension. Furthermore, if predictive inference processes are comparable between the two modalities, we expect to find correlations between the eye-tracking and ERP measures of inference making. Based on earlier findings discussed above, we predict that participants will have difficulty in successfully generating predictive inferences in the low predictability condition compared to the high predictability condition during both reading and listening comprehension. During reading, the relative difficulty of generating inferences will be reflected by longer gaze durations, and more regressions. During listening, relative difficulty in generating inferences will be reflected by a more pronounced N400. We also expect that inference processes will be affected by the WM of participants. On the one hand, if high-WM participants but not low-WM participants generate predictive inferences, then only the high WM comprehenders will show the predictability effect in form of longer gaze durations, more regressions, and a more pronounced N400. On the other hand, if both groups generate predictive inferences but only low-WM participants have difficulty in resolving faulty predictions, then only low-WM participants will show the predictability effect as reflected by of longer gaze durations, more regressions, and a more pronounced N400, in the low-predictive condition compared to the high-predictive condition. If WM plays a greater role in listening than in reading comprehension, these differences would be more pronounced during the listening task than during the reading task. By using a within-subject design, we assess the effect of WM on comprehension processes in the two modalities in the same participants.

5.2 Methods

5.2.1 Participants

Forty-four students at Leiden University participated in the experiment (38 women; average age: 22 years, SD = 2.5 years). All participants were right-handed native Dutch speakers and had normal or corrected-to-normal vision and hearing. They gave written informed consent prior to participating in the study and received a small financial reward or course credits for their participation. Due to technical problems, the data of two participants had to be excluded from the analyses. Data of the remaining 42 participants (36 women; age: 22 years, SD = 2.5 years) was included into the analyses.

5.2.2 Materials and Design

5.2.2.1 Reading and Listening Task

The materials for the listening (ERP) and reading tasks (eye-tracking) were set up in the same way. For both tasks, target words were embedded in short stories consisting of two sentences: a context sentence followed by a target sentence. The second sentence contained the target word, which was a noun with an average length of 6.5 characters (SD = 2.6). The average length of the short stories was 24.3 words (SD = 4.9).

Short stories were designed in such a way that a target word was preceded by an indefinite article and always appeared towards the end of the second sentence (but was never the final word of the sentence). Depending on discourse context, the target word was either high or low in predictability. Each target word appeared in both conditions. The short stories in both high-/low- predictability conditions were semantically correct (e.g., see examples in Introduction). To pre-test the predictability of target words we used a cloze test, which is assumed to be an independent measure of critical word predictability (e.g, Nieuwland, 2015; Nieuwland, 2016). In the cloze task, 20 participants from a similar population as the main sample read the short stories up to the target word and were asked to complete the sentences with the first word that came to their mind. The results of the cloze task showed that participants spontaneously used the high predictability target word (e.g., aquarium) with average cloze probability of 70% (SD = 5.3%).

After having read or listened to each story, participants were presented with a multiple-choice comprehension question with three answer possibilities. The purpose of the questions was to make sure that participants paid attention to the stories. The entire set of materials consisted of 336 short stories. Four stimulus lists were created to counterbalance target words across modality (i.e., reading and listening), with 4 practice items and 80 test sentence pairs and target words per list. Each target word occurred in the two conditions (i.e., high, and low predictability) on different lists, and each participant was exposed to each target word only once.

During the reading task, sentences were displayed 65 cm from the participants' eyes and appeared in Arial font 18 pt size. One degree of visual angle equaled approximately 3.0 characters. Each trial began with a fixation point presented about two character spaces to the left of the first character of the upcoming sentence. The fixation point also served as drift correction, to check whether the calibration was still acceptable. The eye tracker was recalibrated whenever the experimenter deemed necessary. The task was self-paced: after reading a sentence, the participant pressed the spacebar to move to the next sentence and, eventually, to the multiple-choice question. Participants were instructed to read for comprehension and answer the questions as accurately as possible.

During the listening task, short stories were presented auditorily through speakers. All short stories and comprehension questions were recorded by a male native Dutch speaker. Each trial started with the fixation point, which remained on the screen for the duration of the story. The short stories were on average 6141 ms long (SD = 1200 ms). After each short story, a comprehension question was presented auditorily simultaneously with the visual presentation of the answer options. Participants were instructed to listen to the stories for comprehension and answer the question as accurately as possible by pressing the corresponding keys on the button-box.

5.2.2.2 Working-Memory Task

The Mental Counters is a computerized working memory task (Huizinga et al., 2006; Larson et al., 1988). In the Mental Counters task, participants were asked to keep track of and update the score of counters in their working memory. There were two (block 1) or three (block 2) independent counters. The counters are horizontal lines, positioned to the left and to the right from the middle of the computer screen. For each trial, the starting score for each counter was 0. Above and below the lines, squares appeared in a random order. Participants were required to add one point to the value of the counter, when a square appeared above it, and to subtract one point when it appeared below the counter. Before each trial began, participants received a criterion value to remember throughout the trial. Participants were instructed to press a button when any counter reached the given criterion value, before the next block appeared. The score for the task was the proportion of correct trials.

5.2.3 Procedure

Participants were tested individually. After signing the informed consent, participants started with the Mental Counters task. The Mental Counters began with a block of practice items, before moving on to the two test blocks. After completing this task, participants were asked to read and listen to the short stories. The order of the reading and listening tasks were counterbalanced across participants. Reading and Listening tasks started with four practice trials. The entire test session took approximately two hours.

5.2.4 Apparatus

Eye-movement recording. Eye movements were recorded with an Eyelink 1000 eye- tracker (SR Research Ltd.; 500 Hz sampling rate). Eye calibration was done at the beginning of the experiment, using a 9-point calibration procedure. The participant's head was kept immobile with the use of a chin and head rest. Viewing was binocular, but only the movements of the dominant eye were recorded.

Electrophysiological recoding. Electroencephalograms (EEG) were recorded using an EEG cap with 32 active electrodes (Ag/AgCl), mounted according to the extended International 10-20 system. The EEG was collected using BioSemi ActiView, and the EEG signal was digitized at a rate of 512 Hz with a band pass filter of DC-128 Hz. All electrodes were offline re-referenced to the two mastoids. Lateral eye movements were measured using a bipolar montage of two electrodes placed on the right and left external canthus. Vertical eye movements were measured using a bipolar montage of two electrodes placed above and below the eyes.

5.2.5 Analyses

5.2.5.1 Eye-tracking Data

In total, 1% of the data were lost due to eye blinks or technical problems. For each target word, we determined first gaze durations (the sum of all fixation durations on a target word prior to moving to another word), total gaze durations (the sum of all fixation duration on a target word), and regression rates (the percentage of backward eye movements to the target region from a succeeding region). In the analyses of total gaze durations, only those trials were included where the target word was fixated during first-pass reading. The data were subjected to repeated measures ANOVA with Predictability (high- vs. low- predictability) as independent variable. WM group (high-WM vs. low-WM) was used as between-subject factor. The working memory groups were created based on the median split of mental counters scores.

5.2.5.2 Electrophysiological Data

Epochs from –200 ms to +800 ms were obtained relative to the onset of each target, including a 200 ms pre-stimulus baseline. To correct for ocular and non-ocular artefacts, epochs with amplitudes above or below 75 µV were rejected. The EEG signal was applied to a high-pass filter of 0.01 Hz/ 24 dB and a low-pass filter of 40 Hz/ 24 dB. Analyses were performed for a post-stimulus time window of 250 – 550 ms. For this time window, we calculated mean amplitude values per participant and per condition.

Amplitudes were submitted to repeated-measures ANOVAs with Predictability (high- vs. low- predictability), and Location (anterior, i.e. F7, F3, FC1, FC5, AF3, FC6. FC2, F4, F8, AF4, Fz vs. posterior, i.e. CP2, P4, PO4, O2, CP1, CP5, P3, PO3, O1, CP6, Pz) as independent variables. This led to a division of electrodes into two areas, which were used to investigate the distribution of the possible effects. WM group (high-WM vs. low-WM) was used as between-subject factor. The groups were the same as in the analysis of the eye-tracking data.

5.2.5.3 Combining Eye-tracking and Electrophysiological Data

To investigate the relation between reading comprehension (eye-tracking) and listening comprehension (ERP data), we computed Pearson's two-tailed correlations between differences scores (low predictable – high predictable) for First Gaze Durations, Total Gaze Durations, Regression Rates, and the N400 effect.

5.3 Results

5.3.1 Reading Comprehension: Eye-Tracking Data

Participants answered 91.8% (SD = 0.03) of the comprehension questions correctly, showing that they read the sentences attentively.

For the first gaze durations there was no significant effect of Predictability (F < 1). Across both conditions, high-WM participants fixated for shorter duration (M = 240 ms; SD = 32 ms) than low-WM participants (M = 243 ms; SD = 36 ms). However, this difference was not significant ($F(1, 37) = 3.75$, $p = .06$). The interaction between Predictability and WM group was not significant ($F(1, 37) = 1.95$, $p = .17$).

Analysis with total gaze durations as dependent variable showed that all participants spent more time reading target words in the less-predictable condition (M = 471 ms; SD = 287 ms) than in the semantically predictable condition (M = 432 ms; SD = 589 ms; $F(1, 37) = 7.19$, $p = .01$). The main effect of WM group and the interaction between Predictability and WM group were not significant ($F < 1$ and $F(1, 37) = 2.22$, $p = 0.15$, respectively).

Participants were more likely to make regressions to the target in the less-predictable condition (M = .37; SD = .18) than in the high-predictable condition (M = .33; SD = .18; $F(1, 37) = 13.78$, $p = .001$). Across both conditions, there was a lower probability that high-WM participants would look back (M = .28; SD = .18) than that low-WM participants would do so (M = .43; SD = 14; $F(1, 37) = 6.79$, $p = .01$). The interaction between Predictability and WM group was not significant (both $F < 1$).

5.3.2 Listening Comprehension: Electrophysiological Data

Participants answered 90.2% (SD = 0.02) of the comprehension questions correctly, showing that they listened to the sentences attentively.

Analysis of the mean amplitudes showed a main effect of Predictability ($F(1, 37) = 10.77$, $p = .002$). This effect was qualified by a significant interaction between Predictability and WM group ($F(1, 37) = 5.63$, $p = .023$). To further investigate the significant interaction, separate ANOVA's were run for high-WM and low-WM participants. For high-WM participants, there was no significant effect of Predictability (low-predictable: M = -0.99μV; SE = 0.55μV; high-predictable: M = -0.61μV; SE = 0.38μV; $F < 1$, see Figure 5.1). By contrast, for low-WM participants there was a significant effect of Predictability ($F(1, 16) = 14.62$, $p = .001$, see Figure 5.2): N400 amplitudes were more negative for the low predictable condition (M = -1.61μV; SE = 0.41 μV) than for the high predictable one (M = 0.77μV; SE = 0.50 μV;).

Neither main effects of WM group and Location were significant nor were other interactions (Predictability x WM group x Location: $F(1, 37) = 1.62$, $p = .21$; all other effects: Fs < 1).

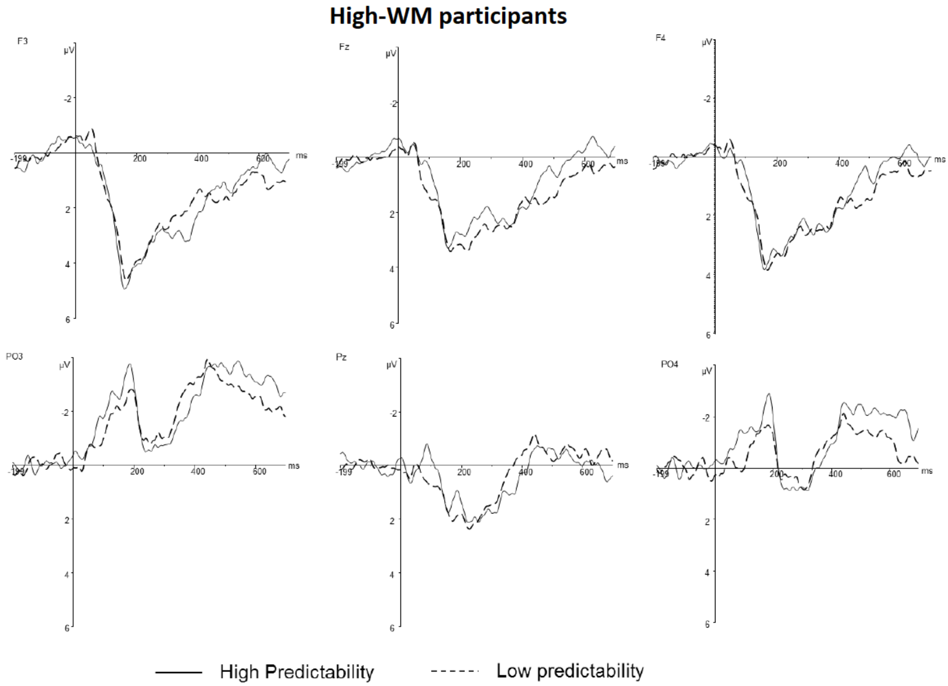

Figure 5.1. Averaged stimulus-locked ERP waveforms for high predictable (solid lines) versus low predictable (dashed lines) target words for high-WM participants. Amplitudes (µV; Y-axis) are plotted against time (ms; X-axis); zero represents onset of the target word.

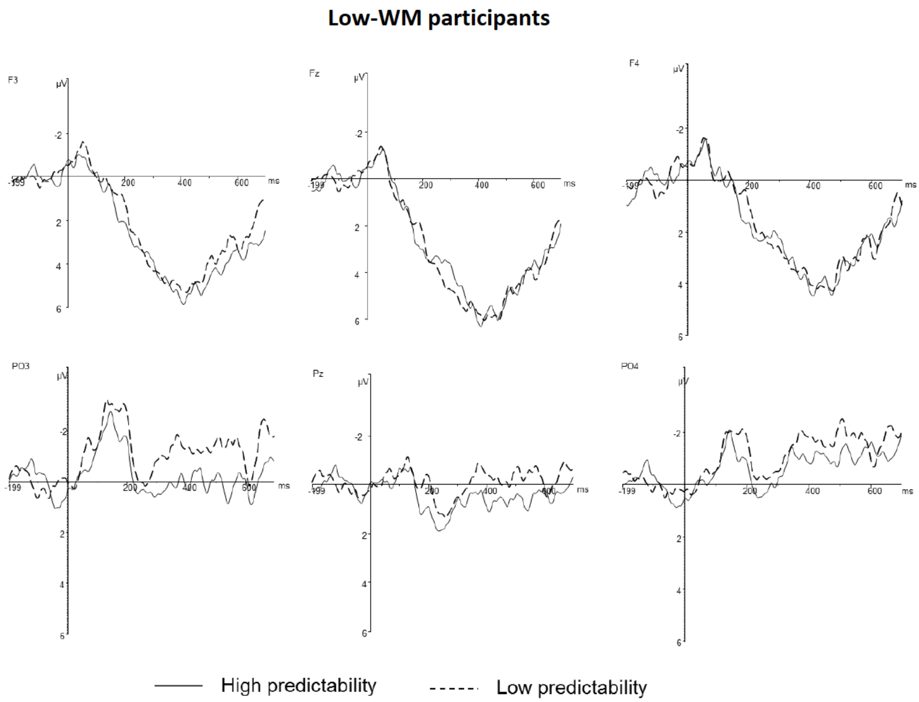

Figure 5.2. Averaged stimulus-locked ERP waveforms for high predictable (solid lines) versus low predictable (dashed lines) target words for low-WM participants. Amplitudes (µV; Y-axis) are plotted against time (ms; X-axis); zero represents onset of the target word.

5.3.3 Relation between Reading and Listening Comprehension

There was a significant positive correlation between first gaze durations and N400 (r = .36; p = .02). Participants who showed larger differences between experimental conditions during reading also had larger differences during listening comprehension. The correlations between total gaze durations and N400, as well as between regression rates and N400 were not significant (r = .15, p = 0.23; r = .12, p = .29, respectively).

5.4 Discussion

The aims of the present study were to investigate (a) whether the role of Working Memory (WM) differs during reading and listening comprehension, and (b) whether the inference generation process is comparable between the two modalities. Overall, for both the reading and listening tasks, we replicated the standard effect. All participants spent more time reading low-predictable words in late reading processes and having more pronounced N400 amplitudes than high-predictable words. However, only during listening this effect was modulated by the WM of participants; low-WM participants, but not high-WM participants, showed an effect of predictability. Furthermore, there was a positive moderate correlation between measures of predictive inference generation during reading and listening comprehension. Below we discuss these findings in more detail.

To investigate the role of WM during reading and listening we recorded eye-movements during reading and EEG during listening in low- and high-WM participants, and found differences between the two language comprehension modalities. Predictive inference processes during reading were not modulated by individual differences in working memory: Participants spent more time reading low-predictable words than high-predictable words. This predictability effect emerged only for total gaze duration, not for first-pass gaze duration, and was due to readers re-fixating on the low-predictable words more often than on the high-predictable words. Because highly predictable words fit better with discourse context they require less overall processing time and, consequently, are easier to integrate into preceding context without regressions. These results are in line with previous findings showing that semantic integration affects late processes more than early processes (e.g., Rayner et al., 2004). Hence, during reading we only see the predictability effect in late processing, and this is independent of WM.

During listening, N400 amplitudes to low-predictable words were more negative than to high predictable ones. This predictability effect was modulated by individual differences in working memory. Low-WM participants, but not high-WM participants, showed more pronounced N400 amplitudes for the low-predictable words compared to the high predictable ones. Similar to our results, Otten and van Berkum (2009) showed a significant predictability effect for low-WM participants but not for high WM participants. Here we propose two possible explanations for the differences in the predictability effect between the two groups: There may be differences in how high- and low-WM participants make predictions, or there may be differences in how high- and low-WM participants process information that is not in line with the made prediction.

First, successful comprehension depends not only on information provided by the discourse context but also on the comprehenders ability to retrieve information from semantic memory (i.e., background knowledge; e.g., Ericsson & Kintsch, 1995; Kintsch, 1998; Till, Mross, & Kintsch, 1988). The N400 reflects not only sensitivity to local context information (e.g., from preceding context) but also the ease with which one accesses information from one's background knowledge (for review see Kutas & Federmeier, 2000). High-WM participants are more efficient in retrieving and/or integrating information from background knowledge than low-WM participants. Consequently, high-WM participants have a wider network of words active in their lexicon during listening (e.g., both bowl and aquarium may be active) thereby decreasing the differences between levels of predictability induced by the local discourse context. Low-WM participants may rely more on recently activated information by the local context than on activation from background knowledge (Boudewyn et al., 2013). This in turn results in a more pronounced difference between levels of predictability induced by manipulating local discourse context. Hence, during prediction generation, high-WM participants use both local context and their background knowledge, whereas low-WM participants rely more on local context alone.

Second, it is possible that the observed differences between high-WM and low-WM participants do not reflect difficulty in making predictions per se but rather a greater effort to suppress their initial prediction about an upcoming target word which was not in line with the local context. Inhibition is an important predictor for successful listening comprehension (Kim & Phillips, 2014) and low-WM participants have more problems with inhibiting unwanted information than high-WM participants do (e.g., Engle, 2002; Gernsbacher & Robertson, 1999). Furthermore, an inability to inhibit irrelevant information may lead to increased N400 amplitudes compared to situations when inhibition was successful (e.g. Debruille, 2007; Debruille et al., 2008; Kutas & Federmeier, 2011). Hence, the N400 effect observed in the low-WM participants may reflect their difficulty suppressing their initial prediction.

To address our second aim, comparing inference generation processes between the comprehension modalities, we compared performance on the reading measures with performance on the listening measure. Results showed that participants who experienced predictive inference difficulties during reading also tended to experience inference difficulties during listening comprehension. The positive correlation was significant only between first gaze duration and N400, and not between total gaze duration or between regression rate and N400. This may be due to the fact that total gaze duration and regression rates reflect later reading processes associated with revisiting the target word, whereas first gaze fixation duration and N400 reflect early comprehension processes. These results lend some support to an amodal view of higher-order comprehension processes. But the correlation is moderate (r = 0.36), leaving enough room for modality-specific differences in processing.

We propose that similar prediction generation mechanisms (e.g., use of local discourse context and background knowledge, and inhibition mechanisms) could play a role during reading as well as listening comprehension. However, a likely modality-specific difference between reading and listening comprehension concerns the demands they make on working memory. During reading there were no differences between high- and low-WM participants concerning the predictability effect. In contrasts, during listening only low- WM participants showed a predictability effect. Spoken language generally does not allow control over the rate of processing and revision of already heard parts of the story (e.g., Constable et al., 2004). Because listening comprehension poses greater demands on the comprehender's cognitive system, a larger WM is of more importance during listening than during reading.

To conclude, whereas most theories of language comprehension take an amodal approach, stating that high order processes are the same between the two modalities, few studies have compared higher-order cognitive processes during reading and listening. Our study expands on previous literature by taking a comparative approach between reading and listening. The results point to specific commonalities as well as differences between the two modalities. Similar underlying mechanisms may play a role during both reading and listening, but our results show that even on a higher-order processing level comprehension during reading and listening may differ. Specifically, we showed that WM plays a greater role in predictive inferencing during listening than reading.

Chapter 6

Summary and General Discussion

6.1 Summary of Results

The empirical studies in this dissertation provide insight in cognitive processes in reading and language comprehension in primarily upper elementary school children. By engaging in different reading processes, readers build a situation model of the text in their mind. The situation model represents events, characters, and situations depicted in the text. We examined how cognitive situation-model building processes are related to reader, text, and task characteristics. Furthermore, we describe the relations among these three factors. Educators know the importance of constructing good learning contexts for students. In their mission to do so, the question "what makes some students succeed and others not?" is an everyday issue. Together with previous research, this dissertation helps in understanding the relations between the three factors and dealing with such questions and the complexity in educational practice and educational science. In this general discussion, I first present a summary of the results from the four empirical papers. Thereafter I discuss the interrelations between reader, text, and task, as well as implications for future research and for education.

In the **second chapter** we studied how children (9-11 years old) differ in online inference generation, and how these differences relate to children's underlying reader characteristics, and text genres. Generating inferences is an elaborated way to build a situation model. Based on think-aloud data, we identified three profiles of readers that differ in both number and types of inferences they generate. Elaborating Readers generated different types of inferences; they used text connecting inferences, elaborative inferences, and predictive inferences while reading. Paraphrasing Readers predominantly repeated the text by paraphrasing it. In addition, they also generated some inferences. Literal Readers mainly repeated the text literally; in fact, 60% of responses from children in this profile were literal text repetitions. Literal Readers generated few inferences. On ancillary measures, Elaborating Readers, showed significantly higher scores on word reading and nonverbal reasoning measures than did Paraphrasing, and Literal Readers. These results show that both lower- order cognitive processes (such as word reading) and higher-order cognitive processes (such as reasoning ability) underpin children's differences in online reading processes. Children generally showed the same reading profiles for both narrative and expository texts. However, generally expository texts elicited fewer elaborative reading processes than narrative texts. For example, there was a smaller number of predictive inferences made when reading expository texts. In addition, a larger number of invalid elaborative inferences were found for expository texts than narrative texts. Because generating fewer elaborative inferences and generating more invalid inferences are likely to impede comprehension, these findings indicate that upper elementary school children are still developing skills to comprehend expository texts. To sum up, the current findings are in line with previous research (e.g. Kraal et al., 2017; McMaster et al., 2012) that has found reader profiles in which children build either a situation model that closely resembles the text or a situation model that is enriched by inferences. In this study, we expand on previous research by showing that such profiles can be found in a group of children with heterogenous underlying cognitive and language abilities, and that these abilities differ between the profiles. In addition, we show that children produce a very similar set of think-aloud responses across text genres that results in the same profiles across the different texts. There were however differences in the number of inferences generated between the text genres. Hence, the results suggest a certain stability in children's text approaches with room for situational differences related to text genre demands.

In the **third chapter** we studied how online processes of children (9-11 years old) in the three reader profiles relate to their offline text memory. Because the offline text memory is an important indicator for learning in school, research of reading profiles needs to consider the quality of resulting memory representations. Offline memory of narrative and expository texts was studied by examining whether children remembered more central information, the gist, than peripheral information, i.e. whether they showed a centrality effect. Inferences help in making connections between text parts, which is important to understand the gist of a text. Therefore, we anticipated that children who generate a larger number of online inferences would show a larger centrality effect in their offline recall, compared to children who generate fewer online inferences. Meaning we hypothesized Elaborating Readers would show a larger centrality effect than Paraphrasing Readers and Literal Readers, and that Paraphrasing Readers would show a larger centrality effect than Literal Readers. First, all groups of children showed a centrality effect for narrative texts. Elaborating Readers showed a larger centrality effect than Paraphrasing Readers. However, neither Elaborating Readers nor Paraphrasing Readers differed from Literal Readers. We suggest that these findings can not be explained only by the number of inferences generated during reading, but we also consider ancillary measures. Elaborating readers proved to achieve higher scores on word decoding and non-verbal reasoning than both Paraphrasing, and Literal Readers. We suggest that children in each profile engage in the online reading processes that suit the cognitive capacities that underpin reading. Although Paraphrasing Readers score lower on word decoding and reasoning tests than Elaborating Readers, Paraphrasing Readers try out some inferences while reading. Possibly, trying out inferences while having relatively underdeveloped word reading and general reasoning abilities hinders Paraphrasing Readers from achieving an offline memory representation similar to that of Elaborating Readers. Second, no centrality effects were found when the children read expository texts. In sum, expository texts elicited fewer inferences in online processes (second chapter) and no profile differences in offline comprehension (third chapter). It is possible that extracting central information in expository texts requires additional online processes to inference generation that our measures did not capture.

In the **fourth chapter** we examined children's (9-12 year old) ability to use temporal connectives when building a situation model of sentences such as "Before you add or subtract a number, you should solve the multiplication". Such sentences can be grammatically complex and therefore taxing for working memory. Previous research has brought contradicting hypotheses of the role of working memory for comprehending these sentences. To expand on previous research, we examine how comprehension interacts with both working memory capacity and working memory updating ability. In two experiments we varied the position of the connectives "before" and "after" in the sentences. In the first experiment we asked participants to answer the question "what happened first?" and in the second experiment we asked participants to answer the question "what happened last?". By these sentence and task manipulations we could investigate whether comprehension was affected by familiarity of the connective, by the position of connective, by the position of the answer (main clause, subordinate clause, or recent clause), and by sentence chronology. Across both experiments, we found that upper-elementary school children's comprehension was affected by clause salience, rather than the familiarity of the connective. The children were sensitive to whether the correct answer to the comprehension question was situated in the main clause or the subordinate clause. Importantly, the second experiment showed that comprehension was qualified by children's working memory updating ability and working memory capacity. Children with high working memory updating showed a main clause advantage, i.e. they performed well when the correct answer was situated in the main clause. Children with low working memory updating showed a sentence-chronology effect i.e their comprehension benefits from chronologically written sentences. The alternative explanation we propose to the sentence-chronology effect, is that children with low working memory updating showed a recency effect, i.e. their comprehension was better when the correct answer was positioned at the end of the sentence. Furthermore, the position of the connective influenced comprehension and, importantly, interacted with working memory abilities. When reading sentences with sentence-initial connectives, readers need to hold information about the connective in working memory for a longer time. For these sentences, children with higher working memory capacity succeeded better in doing so. When reading sentences with sentence-medial connectives, readers need to update their mental representation mid-sentence. For these sentences, children with higher working memory updating ability succeeded better in doing so. Together, these findings indicate that upper-elementary school children's comprehension of sentences containing temporal connectives fits with a working memory capacity-constrained framework of reading comprehension where there are dissociable contributions of working memory capacity and working memory updating.

In the **fifth chapter** we examined comprehension of sentences with target words of high and low predictability in two modalities: reading and listening. Comparing literature on comprehension in pre-school and elementary school children can be difficult as the former may use listening tasks whereas the latter may use reading tasks. Therefore, we examine higher-order cognitive processes in both modalities to better understand differences and similarities between the two comprehension tasks. We started this examination in an adult population. Using ERP, eye-tracking, and a working memory updating task, we investigated whether working memory is taxed differently when reading or listening to sentences leading up to a highly or moderately predictable word. We predicted that a reading task taxes working memory less than a listening task as the reader has the possibility to go back and reread target words that seem less predictable. Indeed, results indicate that working memory is more taxed in the listening task than in the reading task, and comprehension is related to both individual differences in working memory, and task demands such as the possibility to revisit the text. In the reading task, the group of participants with high working memory and the group of participants with low working memory looked equally long at highly and moderately predictable words at first gaze. In addition, both groups of participants looked back at less predictable words equally often. In the listening task, only participants with low working memory showed a pronounced N400 effect towards less predictable words. These results have two alternative explanations. First, participants with low working memory may have difficulties to use both global and local context to predict the target words and focus only on the local context while listening. Participants with high working memory use both global and local context while listening. Second, participants with low working memory have difficulties inhibiting their initial prediction, whereas participants with high working memory area able to inhibit their initial prediction while listening. Because we used a working memory updating task it is more plausible to assume that participants with low working memory updating may have difficulties to use both global and local context. I.e., they are less able to update their understanding depending on the context. However, with these data it is unclear exactly what aspects of working memory updating are concerned (Ecker et al., 2010), or how they were used qualitatively. Further research should clarify these aspects to bring further clarity to the interpretations of results. Finally, there was a moderate positive correlation between the first gaze and N400 measures. This finding indicates that higher-order reading processes, such as predictive inference generation at word level, have commonalities between modalities and are also modality-specific. For the latter, different working-memory demands are made on predictive inferences during reading and listening. Further research should include elementary school children to disentangle conclusions in previous research based on the use of different research methods. Because children's working memory is still developing, it is difficult to make direct comparisons of results concerning an adult population. In addition, research concerning modality differences in children has educational benefits as listening devices are sometimes used in schools as reading aids.

6.2 Interactions Between Reader, Text, and Task Characteristics

As has been pointed out in the introduction to this dissertation and made clear from the results presented in the four empirical chapters, reading comprehension and reading comprehension processes are dependent on at least three sets of factors; namely characteristics of the reader, the text, and the task (see the Venn diagram in Figure 6.1; modified from Snow & RAND, 2002). In this part of the dissertation, I discuss how the four chapters relate to each other; how results indicate an intricate relation between the three factors and how these results could lead to further research and have educational implications. I concentrate on the intersections between the three factors, and pinpoint how our findings contribute to understanding these intersections further. In this discussion, I begin with results related to reader characteristics, reader profiles and working memory, and end with suggestions for further research. In addition, I discuss how the results have educational implications.

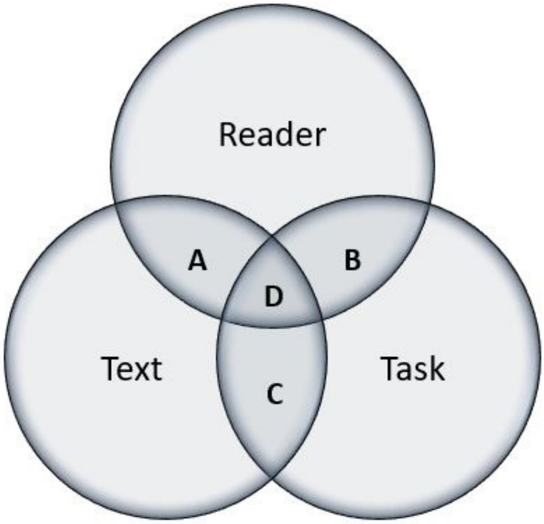

Figure 6.1. Venn diagram depicting three important factors that influence reading comprehension and their interactions: reader, text, and task characteristics.

6.2.1 Reader Profiles

Learning depends on a multidimensional set of skills, and classrooms typically contain a heterogeneous group of students (Hickendorff, Edelsbrunner, McMullen, Schneider, & Trezisee, 2018). Similarly, learning to read requires a multidimensional set of skills and developing readers form a heterogeneous group. Because practical limitations in a classroom context may prevent instruction that is specific to each student, it can be helpful to consider subgroups of developing readers that are similar in some important respects. For example, by identifying reader profiles based on the number and types of inferences that developing readers generate, we were able to make visible relatively homogenous subgroups of elementary-school children and understand how they behave qualitatively different when reading. In addition, by exploring how these subgroups made use of online reading processes and build their offline mental representations in different text genres (field D, figure 6.1) we detected both differences and similarities between these subgroups. Extracting such differences and similarities extends our knowledge on states and traits in elementary-school children's reading abilities. Below I will further discuss these states and traits, and what this implies for future research.

Studying differences in the number of inferences children produce when reading and what they remember from these texts, helps to better understand how online reading processes relate to offline text representation. Together, the empirical papers based on think-aloud data (chapter two) and recall data (chapter three) help us understand this relation. The results from the think-aloud study clarify some aspects of the intersection of reader characteristics and text characteristics (field A, figure 6.1). For example, the appearance of similar reading profiles across two text genres shows that developing readers display a certain stability of reading behavior across different texts and could point to reader traits. The children in the three profiles consistently displayed a similar approach to all texts. Literal Readers stayed very close to the literal meaning of all texts, Paraphrasing Readers took a further step away from the texts by making paraphrases, and Elaborating Readers enriched their situation model of texts furthermore by focusing on inferences. In addition, we found that well-developed skills in both lower-order (word reading ability) and higher-order (reasoning ability) cognitive processes are underpinning the ability to generate inferences in both narrative and expository texts. These findings support the notion of stability in reading behavior displayed across text genres. Although we expand on previous studies by examining these profiles across text genres, the findings resemble previous research as children differ in how close they stay to the literal text and how many inferences they make (e.g. Kraal et al., 2017; McMaster et al., 2012). However, all children responded to a more difficult text genre in a similar way. Expository texts generally introduce new words and conceps to the reader compared to narrative texts and are often considered more difficult (Best et al., 2008; Britton & Pelligrini, 1990). The expository texts elicited fewer inferences for all children. The fact that young readers (have to) partly change their online reading approach in the face of different text demands point to situational states and flexibility. Together, these results indicate both stability and situational states, or in other words continuity and change, in use of inferences across different text genres.

The results from the recall study clarify some aspects at the intersection of reader characteristics and text characteristics as well as at the intersection of reader characteristics and task characteristics (field A and B, figure 6.1). Children in all three profiles remembered more central information than peripheral information after having read narrative texts, i.e. they showed a centrality effect. However, they did not show a centrality effect for expository texts. Possibly, inferences help in building a text representation focusing on central aspects of narrative texts, but other processes may be needed for remembering central aspects of expository texts (Lorch, 2017). The fact that children in all reading profiles displayed a centrality effect for narrative texts but not for expository texts shows similarities in the ability to recall central text elements across reading profiles. Measures from ancillary tasks allow us to speculate why there are large differences between the profiles in the online reading task but not the offline memory representation. In line with previous research (Jonsson et al., 2014; Turley-Ames & Whitfield, 2003), I suggest that children make use of the online reading processes that suit their cognitive abilities. Literal and Paraphrasing Readers showed lower scores on both lower-order and higher-order cognitive abilities compared to Elaborating Readers. These cognitive abilities underpin the ability to make inferences during reading (Carlson et al., 2014; de Leeuw et al., 2016; Naglieri, 2001; Olson, 1985; Rapp et al., 2007). However, for these age-appropriate texts children in all groups are able to create a similar offline text representation. Although similar, there was an important difference between Elaborating Readers and Paraphrasing Readers. Elaborating Readers showed a larger centrality effect than Paraphrasing Readers. These findings could indicate that Paraphrasing Readers are developing or testing inferential online processes that do not yet entirely match their lower and higher order cognitive skills underpinning inference generation. In this, we may call it, exploratory phase, Paraphrasing Readers' online reading efforts are not enough to remember central information from a text in the same way as Elaborating Readers do. Instead, when Paraphrasing Readers use inferences that stretch beyond the base of other cognitive abilities, this may give rise to a somewhat uncertain offline text representation. To summarize, being able to compare performance on different text genres (narrative and expository) and different tasks (online and offline reading measures, and ancillary measures of cognitive capacities) make similarities and differences between the reader profiles visible. In addition, a difference between Elaborating and Paraphrasing Readers in text recall suggests that Paraphrasing Readers use inferences perhaps exploratorily. This raises new research questions on how continuity and change across contextual demands develop over time. Below I will further discuss these questions.

Contrary to our predictions, Literal Readers showed similarities to readers with the other reading profiles in the recall task. However, as expected, Elaborating and Paraphrasing Readers differed from one another. These results raise the question whether different reader profiles depict developmental phases. There is evidence of both stable continuous reading traits and change through contextual demands and development. On the one hand, research of reading profiles based on inferential processes during think-aloud tasks may be pointing to a fairly stable way to describe inter-individual differences. In this study, stability has been shown when children generally kept their profile identity across text genres. In addition, reading profiles in elementary school children have occurred in several studies. Reading profiles have been studied as a means to understand individual differences in 7-8 year olds (Kraal et al., 2017), inference assessment in 8-11-year olds (Carlson et al., 2014), and targeted interventions in 9-10-year olds (McMaster et al., 2012). In addition, a study of university students indicates similar reading profiles exist later in life where groups of readers differ in their use of constructive inferences while reading (Kopatich & Santuzzi, 2018). On the other hand, research of reading development shows a continuously growing ability to use inferences with age, which points to intra-individual changes (Ackerman, 1986; 1988; Kendeou et al., 2008; van den Broek & Kendeou, 2017). In addition, a difference between Elaborating and Paraphrasing Readers in text recall in this study suggests that Paraphrasing Readers use inferences perhaps exploratorily. Therefore, one can speculate whether different reader profiles possibly depict developmental phases. An interesting approach could be to examine the development of these reader profiles over time in longitudinal studies in which we examine intra-individual differences over time in addition to the inter-individual differences.

To sum up this section, extracting differences and similarities across reader profiles, across text genres, and across tasks extends our knowledge on states and traits in elementary-school children's reading abilities. Generally, children in the three profiles used the same approach to text consistently across two text genres. These findings support the notion of stability in reading behavior displayed across text genres. However, the expository texts elicited fewer inferences for all children. The fact that young readers partly change their online reading approach in the face of different text demands point to situational states. Children in all profiles showed a centrality effect in their recall of narrative texts, but not for expository texts. Although children in all profiles resembled each other, there was an important difference between Elaborating and Paraphrasing Readers. Paraphrasing Readers' online reading efforts are not enough to remember central information from a text as well as Elaborating Readers do. This could suggest that Paraphrasing Readers use inferences in an exploratory way that stretches beyond the base of other cognitive abilities and may give rise to a somewhat uncertain offline text representation. Whereas some results point to stable reader traits, other results show change in the face of text demands. Taken together these results call for future studies that are longitudinal to map out continuity and change in these types of reading profiles.

6.2.2 The Role of Working Memory

Working Memory (WM) is seen as an important mental workspace for many cognitive processes, such as reading, writing, and math (Diamond, 2013). However, not everyone would agree on the importance of studying working memory. There have been debates on what working memory actually contains (Cowan, 2017; Diamond, 2013; Miyake & Shah, 1999), and some question whether it is necessary to divide memory in several functions and definitions (e.g. Säljö, 2015), and some empirical data suggest that there may not be a strong relation between working memory capacity and reading (e.g. LARCC & Logan, 2017). However, the results reported in this dissertation show that when we consider under which circumstances working memory plays a role in reading and language comprehension, we can better understand the influence of reader, text, and task. For example, considering the study of children's reading comprehension of temporal relations (chapter four), we found that two aspects of working memory interacted with different textual features. Working memory capacity (holding information in WM while receiving a secondary task) and working memory updating (continuously updating changing information in WM) made separate contributions to children's comprehension of temporal connectives in sentence-initial and sentence-medial position. In addition, children's reading comprehension was modified by individual differences in working memory updating. Furthermore, in adult participants (chapter five), individual differences in working memory updating interact with task demands. In a reading task with relatively low task demands comprehension was not modified by working memory. However, in a listening task with higher task demands working memory modulated adults' comprehension. Hence, studying working memory in relation to variations in contextual demands, such as text and task demands, helps to understand interactions between reader, text, and task.

Aspects from the intersection of reader and text characteristics (field A, figure 6.1) can be made more comprehensible by the results from the study including a population of upper-elementary school children in chapter four. When studying comprehension processes concerning the temporal order of events, clause complexity proved to be a factor that had a negative impact on upper-elementary school children´s reading comprehension. These results show that children of this age do not fully comprehend temporal relations in text. Importantly, the complexity of the text and the children´s understanding for the text interact with their working memory. This interaction shows for both capacity limitations and updating processes in children's working memory. As proposed in the introduction of this dissertation, previous studies of young populations seldomly examine the role of both these aspects of working memory in relation to reading, although models on reading comprehension based on adult populations account for both (see for example McNamara & Magliano, 2009). Our results show that both working memory updating and working memory capacity are positively related to comprehension. Hence, high scores in either working memory test were related to good reading comprehension. Furthermore, distinctions were made between working memory updating and working memory capacity showing that they make separate contributions to the demands of various sentence constructions. When reading sentences making demands on processes necessary to update the content, children with better working memory updating ability performed better, no differences were seen for children's limitations in capacity. And reversed, when reading sentences making demands on a limited-capacity space that holds information in a heightened state of availability, children with better working memory capacity performed better, no differences were seen for updating ability. Together, these results support the idea that models of reading comprehension that include both a capacity limitation and updating processes can be applied on developing populations. Hence, studying interactions between text characteristics, the children's text comprehension, and their working memory makes our theoretical models more comprehensible.

Above I suggest that studying working memory updating, in addition to working memory capacity, is promising for elaborating on reading models in developing populations. This suggestion is strengthened when considering results showing that children of varying working memory updating ability focus their attention on different parts of sentences (chapter four). Children with relatively low working memory updating ability were able to point out the correct answer when it was situated at the end of the sentence. A limited ability to keep and update information in working memory is likely to bias these children to focus on local understanding of sentences, i.e. the last clause read before answering the comprehension question (a so called recency effect; e.g. Naveh-Benjamin, Moscovitch, & Roediger, 2001). Children with a relatively higher working memory updating ability, however, used the main clause as a cue to understand the main message. A good ability to keep and update information in working memory is likely helpful to see past what is most recently read and be able to focus on what is most important. Although this approach did not always render the correct solution, it shows that a larger working memory updating ability helps approach these sentences on a more global level. Gaining understanding for global coherence is important for a reader to create a situation model that includes relations between text parts (van den Broek, Espin, McMaster, & Helder, 2017). Therefore, examining individual differences in understanding global coherence by studying working memory updating makes important contributions to reading research in young populations.

The study including adult readers (chapter five) helps clarify the interaction between reader characteristics and text and task demands (field D, figure 6.1). In addition, it enbles to further elaborate on above mentioned results regarding comprehension of local and global coherence. While reading highly or moderately predictable sentences, participants could control at what speed and how long they engaged in these sentences by, for example, rereading. In this reading situation, there were no differences in reading behavior between participants with low and high working memory updating. However, while *listening* to the same type of sentences they had no possibility to control the speed with which they received the message and no possibility to listen again. Thus, the listening task meant higher demands on working memory and resulted in differences between the two groups. Participants with relatively low working memory updating used the same standards to predict the different types of sentences throughout the listening task, resulting in an N400 response for sentences with lower predictability. Their comprehension response suggests that their relatively low working memory limited their ability to understand the task as a whole. Instead, they were biased towards understanding each sentence separately, i.e. understanding local coherence. By contrast, adults with relatively high working memory could comprehend global coherence. Possibly, a higher working memory allowed them to understand the task as a whole and flexibly interpret sentences of higher or lower predictability. Hence, task demands interact with individual differences in working memory and textual demands. These findings warrant further research on working memory updating and its interaction with task demands in adults. Examining whether adults adopt a certain strategy towards task demands as the task is progressing, and whether potential learning effects differ between adults with high and low working memory, could possibly strengthen models on reading comprehension aiming to explain strategic, reader-initiated, processes. A possible hypothesis is that readers with low working memory updating would be biased towards focusing on local coherence throughout the task, whereas readers with high working memory updating would eventually evolve an understanding for the task as a whole and start using more strategic processes aiming for a global understanding of the task and its demands.

As there is potential for further research in adults concerning the impact of task demands on working memory and comprehension, I also suggest new research questions concerning task effects on developing readers' working memory and comprehension. Because children in modern schools can receive digital reading aids that read the text for the student (Magnusson Amu, 2020), it is important to study to which degree and why children's language comprehension could differ between reading tasks and listening tasks. To know when and how to employ such read-aloud aids optimally, research that examines both reading and listening modalities is necessary. In the current dissertation, we found that both adults and children with high working memory updating seem able to focus on a global understanding of task and sentences. However, individuals with relatively low working memory updating seem biased towards local coherence, especially in the face of higher task or text demands. Therefore, I hypothesize that a listening task will be more demanding for children with low working memory compared to children with high working memory. Hence, if listening tasks are used as reading aids in school it is important to research possible demands they impose on children to know when and how to use these aids. Educators need to monitor differences between various reading devices to understand when they are helpful, and when they impose new challenges. For example, if children with low working memory are prone to focus on local text coherence, they need a reading or a listening context that helps them to also understand the global coherence of the text.

To summarize, the results reported in this dissertation show that when we consider under which circumstances working memory plays a role in reading and language comprehension, we can better understand the interaction of reader, text, and task. Whereas previous research in young populations often focused on working memory capacity, we tested both working memory capacity and updating. The results show that working memory capacity and updating make separate contributions to explain children's reading comprehension. Therefore, models on reading comprehension that include both a limited-capacity space holding information in a heightened state of availability and processes necessary to update the contents of this space can be applied to developing populations. Furthermore, individual differences in working memory are visible in both adults and children, the results presented in this dissertation indicate that working memory in interaction with both task and text demands modifies comprehension and to which parts attention is allocated. Children and adults with relatively low working memory updating are biased towards comprehension of local coherence. Children and adults with relatively high working memory updating could comprehend several parts of the task or sentences, which shows comprehension of global coherence. Based on these findings, I have suggested further research in adults and children that may help build our theoretical knowledge on the relation between working memory and global and local coherence and may have implications for educational practices.

6.3 Educational Implications

Reader characteristics, such as the ability to generate inferences and to store and use information in working memory, can best be explained when considered in relation to contextual demands, such as text and task characteristics. If scientific methods can provide these explanations, they are useful to educational practice. Learners are a heterogeneous group and there are no 'one size fits all' solutions in educational situations. Therefore, scientific explanations need to consider the complexity of reading comprehension. Specifically, the results reported in this dissertation show under which circumstances inference generation and working memory are limited or enhanced in reading and language comprehension. By highlighting two educational implications from these results, this discussion aims to assist educators to create a suitable learning environment for students with various cognitive functioning. In both examples, using the Venn diagram is suggested as helpful because it visualizes three two-way interactions as well as the three-way interaction of reader, text, and task characteristics.

The first educational implication considers recognizing students' varying abilities to build a coherent mental model of a text under different circumstances. To understand and learn from texts, readers have to construct a coherent situation model of the text in their mind. For their situation model to be coherent, it needs to contain information elements from the text, and, importantly, semantic relations between these elements (e.g. van den Broek, et al., 2017). Although important, it is not enough to understand single sentences of a text, i.e. local coherence. Obtaining comprehension of global coherence is necessary to understand the meaning of the whole text. However, there are individual differences in how school children obtain global coherence. In the current dissertation, several results indicate differences in the ablity to obtain global coherence. For example, in the think-aloud study (chapter two), Elaborating Readers use text-connecting inferences, elaborate inferences, and predictive inferences as global coherence-building processes during reading. Paraphrasing Readers made fewer inferences than Elaborating Readers, and Literal Readers made fewer inferences than both previous groups during reading. In addition, when meeting expository texts, all children made fewer inferences. In the two later studies where we compare readers of varying working memory abilities (chapter four and five), both adult and young readers with high working memory were biased towards global coherence. However, adult and young readers with low working memory were biased towards local coherence, especially in a more demanding context. Hence, to foster development of comprehension processes it is important that educators are knowledgeable about and can create circumstances under which students easily learn and access both local and global coherence processes. To do so, educators could use the Venn diagram presented in the beginning of this chapter. By considering the interrelation between the three elements of reading comprehension, the reader, the text, and the task, educators can account for several spects when assessing what is needed for certain reading goals in different situations and for various students. The challenge of text and task can be chosen both in relation to one another and in relation to the students' abilities and set the students up for successful learning.

The second educational implication considers recognizing similarities and differences between teaching material and assessment material, and in doing so accounting for text and task demands. The studies of the think-aloud reading profiles may form an example. In the think-aloud study (chapter two), expository texts impede inference generation in all reading profiles, compared to narrative texts for which children could generate more inferences. In addition, in the offline task, children were not able to extract central information from expository text (chapter three). Together these results concur with previous research concluding that text genre influences comprehension and that expository texts are more difficult for developing readers compared to narrative texts (Best et al., 2008; Eason et al., 2012). Furthermore, whereas the online task seemed sensitive enough to assess individual differences in reading comprehension processes, the offline task seemed less sensitive. When combining the text and task perspective on reading comprehension processes, we see an interesting interaction. In the offline task investigating expository texts, there were no individual differences at all. Possibly the impeding effect of expository texts and the lower sensitivity of the offline task decreases the chances to assess and understand students' abilities. In an educational context, this has implications for various goals of teaching and assessing reading comprehension. For example, when teaching skills that are important for comprehension of a certain type of text (Lorch, 2017), the same text genre needs to be used in the assessment as well. If, however, the goal is to assess students' ability to transfer skills and apply their knowledge in other contexts, there is a need to use assessment material that is differing from the teaching material. In this latter situation, it is important to possess knowledge of the effect of that assessment material. I.e. it is important to understand the effects of text and task characteristics, to understand what skills the students actually transfer to another setting. For example, if teaching online inference generation in narrative text and assessing inference generation in another task or text genre, one must be aware that it is not only the students' abilities that are being tested, but that the task and text also provide certain restrictions for the students to apply their knowledge. Hence, learning can be evaluated by using both similar and differing material. However, the assessor needs to be sure of what the aim is, e.g. context-dependent or context-independent (transfer) learning, and the limitations of the material. Using the Venn diagram as a guide to choose assessment tools could be helpful in both situations. Knowledge about the three elements, reader, text, and task, and taking into account the possible ways they interact can help estimating what to expect from the assessment situation. In this sense, the Venn diagram ensures to always consider how students with varying abilities react to teaching and assessment methods of varying characteristics to get a more complete picture of their learning possibilities.

6.4 Conclusion

In this dissertation, I report research on interacting effects of reader, text, and task characteristics in 9-12 year old children. Within this three-way interaction I have focused on how situation-model building cognitive processes differ in various reading profiles and in readers with low and high working memory. By defining subgroups of readers, either by the inferences they generate or cognitive capacities such as working memory, we gain understanding in how multifaceted and complex reading comprehension development is. In a heterogeneous group of children we found three reader profiles that use different approaches when trying to comprehend the text. Literal, Paraphrasing, and Elaborative Readers vary in the number of inferences they generate while reading. By investigating how these subgroups interact with text demands we understand that children keep their approach and thus behave similarly although facing different text genre demands. However, comparing the reader profiles on another task rendered more similarities between the groups than differences. Thereby, we can generalize profiles over text genres but not across tasks.

Research on individual differences and subgroups based on working memory capacity and working memory updating show that comprehension is modulated by working memory, especially in a context with higher demands. Differing hypotheses from previous research could be solved by examining both working memory capacity and working memory updating in young readers. Working memory capacity and working memory updating make different contributions to reading comprehension in interaction with textual factors. In addition, both adult and young readers are biased towards local or global comprehension depending on having low or high working memory updating abilities, respectively. Thereby, reading models considering both capacity limitations and updating abilities apply also to elementary-school children.

Finally, all four empirical papers show that subgroups differ in whether they focus predominantly on local coherence or try to comprehend global coherence of text. In addition to understanding local coherence, understanding global coherence allows the reader to build a more interconnected situation model of the text. Considering that both reader characteristics and contextual demands enable and limit abilities to comprehend global coherence, educational practices and research need to keep these elements in mind.

Nederlandse samenvatting

Begrijpend lezen bij kinderen op de basisschool:

Cognitieve studies van de lezer, de tekst en de taak

Begrijpend lezen is een veelzijdige set van vaardigheden die essentieel is om deel te nemen aan de moderne maatschappij; bijvoorbeeld om te leren op school, voor werkgerelateerde communicatie, voor sociale digitale interactie en om op de hoogte te blijven van het nieuws. Belangrijke veranderingen in de ontwikkeling van deze vaardigheden vinden plaats tussen de leeftijd van 9 en 12 jaar, wanneer kinderen op de basisschool overgaan van 'leren lezen' naar 'lezen om te leren'. In deze fase verwachten leerkrachten dat kinderen kun leesvaardigheden inzetten om kennis over verschillende onderwerpen te vergaren. Kinderen staan echter niet alleen in hun reis om vaardige lezers te worden. Onderzoek en onderwijsmiddelen worden ingezet om hen op hun pad te ondersteunen. De vier empirische hoofdstukken in het huidige proefschrift presenteren onderzoek vanuit een cognitief-wetenschappelijk perspectief, gericht op drie elementen die belangrijk zijn om meer inzicht te krijgen in begrijpend lezen: de lezer, de tekst, en de taak (Snow & RAND, 2002; van den Broek, Fletcher, & Risden, 1993).

In de cognitieve wetenschap wordt het verkrijgen van diep tekstbegrip beschreven als de constructie van een mentaal model ofwel een situatiemodel (Johnson-Laird, 1983; van Dijk & Kintsch, 1983). De beschreven situaties, gebeurtenissen, en personen moeten in het hoofd van de lezer worden voorgesteld. Het construeren van een situatiemodel hangt af van kenmerken van de lezer, de tekst, en de taak. Ten eerste zijn diverse cognitieve vaardigheden en strategieën nodig om een situatiemodel te construeren, zoals het maken van inferenties tijdens het lezen van de tekst (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Graesser, Kintsch, 1994; Singer, & Trabasso, 1994). Bij het genereren van inferenties, ofwel het 'tussen de regels lezen', legt de lezer verbindingen tussen verschillende delen van de tekst onderling en tussen delen van de tekst en haar voorkennis. Door in te zien hoe zinnen onderling samenhangen en het situatiemodel te verrijken met haar voorkennis, begrijpt de lezer dat de tekst niet slechts een rijtje woorden is, maar dat deze een verhaal vertelt met ontwikkelingen in gebeurtenissen, met oorzaken en gevolgen. Processen of strategieën, zoals inferenties die tijdens het lezen worden gemaakt, worden *online* leesprocessen genoemd. Deze online processen dragen bij aan de herinnering van de tekst na het lezen. De herinnering van de tekst na het lezen wordt een *offline* herinnering genoemd. Ten tweede hebben verschillen in genre, onderwerp, moeilijkheidsgraad, en lengte van de tekst een effect op wat de lezer uit de tekst kan extraheren, en daarmee ook op hoe rijk het situatiemodel dat de lezer construeert kan worden (bv. van den Broek et al., 1993). De lezer kan gemakkelijk een rijk situatiemodel maken bij het lezen van een makkelijke tekst, zoals een tekst met een duidelijke structuur en alledaags taalgebruik, en bereikt zo een goed begrip van de inhoud. Zodra de tekst echter uitdagender wordt, zoals door het introduceren van nieuwe woorden en concepten, moet de lezer harder werken om tot een goed begrip te komen. Ten derde heeft de taak die de lezer krijgt opgelegd, of zichzelf oplegt, tijdens het lezen een effect op het vermogen om een rijk situatiemodel te maken (bv. van den Broek & Kendeou, 2017). Het op eigen tempo lezen van een tekst met de mogelijkheid om zinnen die lastig te begrijpen zijn te herlezen, zal tot een andere mate van begrip leiden dan wanneer dezelfde tekst wordt gelezen zonder controle te hebben over het tempo waarop de tekst wordt gepresenteerd, zoals bij voorlezen of bij een digitaal hardopleesprogramma. De vele complexe cognitieve processen, zoals inferenties, die bijdragen aan begrijpend lezen hebben dus een interactie met kenmerken van de lezer, de tekst en de taak (Rapp & van den Broek, 2005).

Het onderzoek in dit proefschrift richt zich met name op het vermogen om een situatiemodel van een tekst te maken van Nederlandse kinderen in de leeftijd van 9-12 jaar. Deze periode is een belangrijke fase in de leesontwikkeling. In de leeftijd van 9-12 jaar vindt in veel Westerse onderwijssystemen zoals het Nederlandse systeem een overgang plaats van 'leren lezen' naar 'lezen om te leren' (bv. Chall 1983; 1996; Poolman, Leseman, Doornenbal, & Minnaert, 2017). In de lezen-om-te-leren fase wordt van de kinderen verwacht dat ze vloeiend woorden kunnen decoderen en krijgen ze langere teksten met als doel om inhoud aan te bieden. Deze teksten vereisen echter aanvullende leesprocessen ten opzichte van de makkelijke teksten die in de eerste fase werden gebruikt. De leesprestaties van kinderen nemen vaak af bij de overgang van de onderbouw naar de bovenbouw van de basisschool. Dit fenomeen geeft aan dat de overgang van leren om te lezen naar lezen om te leren niet geleidelijk verloopt en onderstreept dat het lezen van ingewikkelder teksten hogere eisen stelt aan de lezer (Chall 1983; 1996). Als onderwijsontwikkelaars leesinstructies zorgvuldig opstellen, kunnen kinderen ook de moeilijke teksten goed begrijpen.

Hoewel kinderen in de Nederlandse bovenbouw te maken krijgen met hogere eisen en mogelijkheden in hun school- en thuisomgeving, is hun cognitieve systeem nog in ontwikkeling. Executieve functies en hersennetwerken die taalvaardigheden en executieve functies ondersteunen blijven zich ontwikkelen bij oudere kinderen en tot ver in de late adolescentie (bv. Diamond, 2013; Gathercole, Pickering, Ambridge, Wearing, 2004; Huizinga, Dolan, & van der Molen, 2006; Zielinski, Gennatas, Zhou, & Seeley, 2010). Executieve functies kunnen omschreven worden als een paraplubegrip voor cognitieve processen, die controle over gedachten en gedrag mogelijk maken (Diamond, 2013), zoals het werkgeheugen. Het werkgeheugen kan worden gezien als een mentale werkruimte die het mogelijk maakt om informatie actief in gedachten te houden en met deze informatie te werken door processen zoals het updaten van de inhoud in het werkgeheugen (Daneman & Merikle, 1996; Diamond, 2013). Het werkgeheugen is belangrijk in de context van begrijpend lezen omdat het ten grondslag ligt aan de capaciteit en processen die nodig zijn om de inhoud van het verhaal op te slaan en bij te werken terwijl de lezer door de tekst gaat (bv. Carretti, Borella, Cornoldi, & De Beni, 2009; Daneman & Merikle, 1996).

De vier empirische studies in dit proefschrift geven inzicht in de cognitieve processen die een rol spelen bij lezen en taalbegrip bij voornamelijk kinderen uit de bovenbouw van het basisonderwijs. We onderzochten hoe cognitieve processen voor het bouwen van situatiemodellen samenhangen met kenmerken van de lezer, de tekst en de taak. Verder beschrijven we de relaties tussen deze drie factoren. Leerkrachten kennen het belang van het creëren van goede leercontexten voor leerlingen. In hun missie om dit te bereiken komt de vraag "wat maakt dat sommige leerlingen slagen en andere niet?" vaak voor. Samen met eerder onderzoek helpt dit proefschrift bij het begrijpen van de relaties tussen de drie factoren en bij het omgaan met de complexiteit van de onderwijspraktijk en onderwijswetenschappen.

In het **tweede hoofdstuk** wordt een onderzoek beschreven naar hoe kinderen (9-11 jaar oud) verschillen in het maken van *online* inferenties en hoe de verschillen hierin samenhangen met achtergrondkenmerken van de lezer. Kinderen lazen teksten uit twee genres: narratieve en informatieve teksten. Op basis van hardopdenk-gegevens hebben we drie profielen van lezers geïdentificeerd die verschillen in zowel het aantal als de soorten inferenties die ze maakten. *Uitweidende lezers* genereerden verschillende soorten inferenties: ze gebruikten tekstverbindende inferenties, uitweidende inferenties en voorspellende inferenties tijdens het lezen. *Parafraserende lezers* herhaalden de tekst voornamelijk door deze te parafraseren, of samen te vatten, terwijl ze ook enkele inferenties maakten. *Letterlijke lezers* herhaalden voornamelijk de tekst letterlijk: Gemiddeld waren zelfs 60% van de reacties van kinderen met dit profiel letterlijke tekstherhalingen. Letterlijke lezers maakten weinig inferenties. Op aanvullende gemeten factoren lieten Uitweidende lezers hogere scores zien op het gebied van woordlezen en non-verbaal redeneren dan Parafraserende en Letterlijke lezers. Deze resultaten laten zien dat wat betreft kenmerken van de lezer zowel lagere-orde cognitieve processen (zoals het lezen van woorden) als hogere-orde cognitieve processen (zoals redeneervermogen) de verschillen tussen kinderen in *online* leesprocessen ondersteunen. Kinderen vertoonden over het algemeen dezelfde leesprofielen voor narratieve als voor informatieve teksten. Over het algemeen leidden informatieve teksten echter wel tot minder uitweidende leesprocessen dan narratieve teksten. Er werd bijvoorbeeld een kleiner aantal voorspellende inferenties gemaakt bij het lezen van informatieve teksten. Bovendien werd een groter aantal foute uitweidende inferenties waargenomen voor informatieve teksten dan voor narratieve teksten. Omdat het genereren van minder uitweidende inferenties en het genereren van meer foute inferenties het begrip waarschijnlijk belemmeren, geven deze bevindingen aan dat de vaardigheden die nodig zijn om informatieve teksten te begrijpen nog steeds in ontwikkeling zijn bij kinderen in de bovenbouw van de basisschool. Samenvattend zijn de huidige bevindingen in overeenstemming met eerdere onderzoek (bv. McMaster et al., 2012; Kraal et al., 2017) waarin lezersprofielen zijn gevonden waarin kinderen ofwel een situatiemodel bouwen dat sterk lijkt op de tekst, ofwel een situatiemodel dat is verrijkt met inferenties. Met deze studie breiden we eerder onderzoek uit door aan te tonen dat dergelijke profielen te vinden zijn in een groep kinderen met heterogene onderliggende cognitieve en taalvaardigheden, en dat deze vaardigheden verschillen tussen de profielen. Bovendien laten we zien dat kinderen een zeer vergelijkbare reeks hardop-denk-antwoorden produceren in verschillende tekstgenres, wat resulteert in dezelfde profielen voor de verschillende teksten. Er waren echter verschillen in het aantal gegenereerde inferenties tussen de tekstgenres. Daarom suggereren de resultaten een zekere stabiliteit in de tekstbenaderingen van kinderen met ruimte voor situationele verschillen die verband houden met eigenschappen van de tekst.

In het **derde hoofdstuk** hebben we onderzocht hoe de *online* processen van kinderen (9-11 jaar) in de drie lezersprofielen (gepresenteerd in hoofdstuk twee) zich verhouden tot hun *offline* tekstgeheugen. Omdat het offline tekstgeheugen een belangrijke indicator is voor leren op school, moet bij onderzoek naar leesprofielen rekening worden gehouden met de kwaliteit van de geheugenrepresentaties. Het offline geheugenrepresentatie van narratieve en informatieve teksten werd onderzocht door in kaart te brengen of kinderen zich na het lezen meer centrale informatie, de kern, dan perifere informatie uit de tekst herinnerden, d.w.z. of ze een centraliteitseffect vertoonden. Inferenties helpen bij het maken van verbindingen tussen tekstdelen, wat belangrijk is om de kern van een tekst te begrijpen. Daarom verwachtten we dat kinderen die een groter aantal online inferenties maken een groter centraliteitseffect zouden laten zien in hun offline herinnering, vergeleken met kinderen die minder online inferenties maken. Dit betekent dat we veronderstelden dat Uitweidende lezers een groter centraliteitseffect zouden laten zien dan Parafraserende lezers en Letterlijke lezers, en dat Parafraserende lezers een groter centraliteitseffect zouden laten zien dan Letterlijke lezers.

Ten eerste vertoonden alle groepen kinderen een centraliteitseffect voor narratieve teksten. Uitweidende lezers lieten een groter centraliteitseffect zien dan Parafraserende lezers. De Uitweidende en de Parafraserende lezers verschilden echter beiden niet van Letterlijke lezers. Een mogelijke verklaring is dat we naast het aantal inferenties dat tijdens het lezen wordt gemaakt ook rekening moeten houden met andere kenmerken van de lezer. Uitweidende lezers bleken hogere scores te behalen op het decoderen van woorden en non-verbaal redeneren dan zowel Parafraserende als Letterlijke lezers. Mogelijk houden kinderen in elk profiel zich bezig met online leesprocessen die passen bij de cognitieve capaciteiten die ten grondslag liggen aan hun lezen. Hoewel Parafraserende lezers lager scoorden op het decoderen van woorden en redeneren dan de Uitweidende lezers probeerden Parafraserende lezers tijdens het lezen wel enkele inferenties uit. Mogelijk is het uitproberen van inferenties belemmerend als woordlees- en algemene redeneervaardigheden relatief onderontwikkeld zijn bij Parafraserende lezers, en daarom lukt het hen niet om een offline geheugenrepresentatie te vormen die lijkt op die van Uitweidende lezers. Ten tweede werden er geen centraliteitseffecten gevonden wanneer de kinderen informatieve teksten lazen. Samenvattend, informatieve teksten lokten minder inferenties uit in online processen (tweede hoofdstuk) en geen profielverschillen in offline begrip (derde hoofdstuk). Het is mogelijk dat het extraheren van centrale informatie uit informatieve teksten aanvullende online inferentieprocessen vereist die onze metingen niet hebben vastgelegd. In eerder onderzoek zijn vaak afzonderlijke inferentieprocessen in korte teksten onderzocht. In het huidige onderzoek hebben we verschillende inferenties onderzocht bij het lezen van langere teksten, en daarmee ontdekten we dat het effect van inferentieprocessen op het offline geheugen niet eenduidig is, maar dat andere eigenschappen van de lezer gecombineerd met het maken van inferenties het offline geheugen voor de tekst mogelijk verklaren.

In het **vierde hoofdstuk** hebben we het vermogen van kinderen (9-12 jaar) onderzocht om temporele verbindingswoorden te gebruiken bij het construeren van een situatiemodel voor zinnen zoals "Voordat je een getal optelt of aftrekt, moet je de vermenigvuldiging uitrekenen". Dergelijke zinnen kunnen grammaticaal complex zijn en daarom belastend voor het werkgeheugen. Eerder onderzoek heeft tegenstrijdige hypothesen opgeleverd over de rol van het werkgeheugen bij het begrijpen van dit soort zinnen. Om dit eerdere onderzoek uit te breiden onderzochten we de interactie tussen begrip enerzijds en de capaciteit van het werkgeheugen en het vermogen om het werkgeheugen te updaten anderzijds. In twee experimenten hebben we de positie van de verbindingswoorden "voordat" en "nadat" in de zinnen gevarieerd. In het eerste experiment stelden we de deelnemers de vraag "wat is er als eerste gebeurd?" en in het tweede experiment stelden we de vraag "wat is er als laatste gebeurd?". Door deze zin- en taakmanipulaties konden we onderzoeken of het begrip werd beïnvloed door de bekendheid van het verbindingswoord, door de positie van het verbindingswoord, door de positie van het antwoord (in de hoofdzin, in de bijzin, of in de meest recente zinsdeel), en door de chronologie van de zin. Bij beide experimenten ontdekten we dat het begrip van kinderen in de bovenbouw van de basisschool meer werd beïnvloed door de betekenis van de zinsdelen, voornamelijk de hoofdzin, dan door de bekendheid van het verbindingswoord. De kinderen waren gevoelig voor of het juiste antwoord op de begripsvraag in de hoofdzin of de bijzin lag. Belangrijk hierbij is dat het tweede experiment aantoonde dat begrip werd bepaald door het vermogen van kinderen om het werkgeheugen te updaten en door hun werkgeheugencapaciteit. Kinderen met een hoog vermogen om het werkgeheugen te updaten waren gevoelig voor de rol van de hoofdzin: ze presteerden goed wanneer het juiste antwoord in de hoofdzin stond. Kinderen met een laag vermogen om het werkgeheugen te updaten vertoonden juist een recentheidseffect: hun begrip was beter wanneer het juiste antwoord aan het einde van de zin stond. Deze resultaten suggereren kwalitatieve verschillen tussen het leesbegrip van kinderen met een laag en een hoog vermogen het werkgeheugen te updaten. Bovendien beïnvloedde de positie van het verbindingswoord het begrip en, belangrijker nog, vertoonde dit een interactie met het werkgeheugen. Bij het lezen van zinnen met het verbindingswoord aan het begin van de zin moeten lezers informatie over het verbindingswoord in het werkgeheugen opslaan totdat ze de volledige zin hebben gelezen, en vervolgens de informatie die het verbindingswoord impliceert toepassen. Kinderen met een hogere werkgeheugencapaciteit slaagden daar beter in. Bij het lezen van zinnen met het verbindingswoord in het midden van de zin moeten lezers hun mentale representatie midden in een zin bijwerken. Daar slaagden kinderen met een hoger vermogen tot het updaten van he werkgeheugen beter in. Samen geven deze bevindingen aan dat het begrip dat kinderen in de bovenbouw hebben van zinnen die temporele verbindingswoorden bevatten wordt beïnvloed door kenmerken van de lezer waarbij de invloed van de capaciteit van het werkgeheugen en het vermogen tot het updaten van het werkgeheugen kunnen worden onderscheiden.

Tegenstrijdige hypothesen in eerdere literatuur over verbindingswoorden kwamen voort uit twee kwesties: ten eerste was dit onderzoek gericht op ofwel de capaciteit van het werkgeheugen ofwel het updaten van het werkgeheugen (resultaten gerelateerd aan kenmerken van de lezer), en ten tweede dit onderzoek gebruikte ofwel een leestaak ofwel een luistertaak (resultaten gerelateerd aan kenmerken van de taak). In hoofdstuk vier hebben we meer duidelijkheid gegeven over de relatie tussen de twee aspecten van het werkgeheugen. Het vijfde hoofdstuk beschrijft onderzoek naar de rol van kenmerken van de taak.

Het **vijfde hoofdstuk** beschrijft een studie waarin we onderzochten hoe zinnen met doelwoorden met een hoge of lagere voorspelbaarheid in twee modaliteiten, lezen en luisteren, werden begrepen. Dit onderzoek beschrijft een eerste studie in een volwassen populatie (gemiddelde leeftijd 22 jaar). De deelnemers lazen zinnen met doelwoorden van hoge en lage voorspelbaarheid. In het volgende voorbeeld heeft het woord ‘aquarium’ een hogere voorspelbaarheid dan het woord ‘kom’: “Peter vindt tropische vissen heel erg mooi. Thuis heeft hij veel verschillende soorten in een aquarium/kom zwemmen”. Terwijl in het volgende voorbeeld het woord ‘kom’ meer voorspelbaar is: “Het jongetje was ontzettend blij met zijn nieuwe goudvis. Thuisgekomen doet hij de vis meteen in een kom/aquarium met schoon water.” We verwachtten dat het begrijpen van woorden met een lage voorspelbaarheid het werkgeheugen meer belast dan het begrijpen van woorden met een hoge voorspelbaarheid. Daarnaast voorspelden we dat een leestaak het werkgeheugen minder belast dan een luistertaak, aangezien de lezer de mogelijkheid heeft om terug te gaan en doelwoorden te herlezen die minder voorspelbaar lijken. Bij het luisteren heeft men niet de mogelijkheid om terug te luisteren. We onderzochten of het updaten van het werkgeheugen anders wordt belast bij het begrijpen van deze zinnen; tijdens het luisteren met behulp van de ERP component N400 (De N400 component van een ERP is een online maat voor het verwerken van de betekenis van woorden tijdens luisteren), en tijdens het lezen met behulp van oogbewegingenonderzoek (de tijd die een lezer op een woord fixeert is een maat voor het verwerken van de betekenis van dat woord). Ten eerste was er een matige positieve relatie tussen metingen in de leestaak en de luistertaak. Deze bevinding geeft aan dat er zowel overeenkomsten als verschillen zijn tussen begripsprocessen in lezen en luisteren. De resultaten laten zien dat het werkgeheugen meer belast wordt bij de luistertaak dan bij de leestaak, en dat begrip zowel verband houdt met lezerskenmerken van werkgeheugen (relatief hoog of laag werkgeheugen) als met taakkenmerken (lezen of luisteren). Bij de leestaak waren er geen verschillen in het lezen tussen deelnemers met een hoog werkgeheugen en een laag werkgeheugen. In de luistertaak vertoonden alleen deelnemers met een laag werkgeheugen een meer uitgesproken N400-effect in reactie op minder voorspelbare woorden, wat aantoont dat ze moeite hadden deze te begrijpen. Daarom is er een wisselwerking in de beïnvloeding van het begrip tussen de kenmerken van het werkgeheugen van de lezer en de taakkenmerken. Onderzoek naar modaliteitsverschillen bij kinderen kan voordelen voor het onderwijs hebben aangezien op scholen soms luistersoftware wordt gebruikt als leeshulp. Verder onderzoek moet zich echter ook richten op basisschoolkinderen omdat het moeilijk is om directe vergelijkingen te maken van de resultaten bij volwassenen, aangezien het werkgeheugen van kinderen nog in ontwikkeling is.

Samenvattend hebben kenmerken van de lezer - zoals het vermogen om inferenties te maken, technisch lezen, redeneervermogen en het werkgeheugen - een wisselwerking met kenmerken van de tekst en van de taak. Om de ontwikkeling van begrijpend lezen bij kinderen volledig te begrijpen, moet bij onderzoek rekening worden gehouden met de kenmerken van de context, zoals de tekst en de taak, waarbinnen het lezen plaatsvindt. Dit houdt in dat leerkrachten de leesontwikkeling van kinderen in elke nieuwe context moeten ondersteunen.

Svensk sammanfattning

Läsförståelse hos grundskolebarn:

Kognitiva studier av läsaren, texten och uppgiften

Läsförståelse är en mångfacetterad färdighet som är viktig att förvärva för att kunna delta i det moderna samhället: för att lära sig i skolan, för arbetsrelaterad kommunikation, för digitala sociala interaktioner och hålla sig uppdaterad med nyheter. Viktiga förändringar och utveckling i läsförståelse uppstår mellan 9 och 12 år när grundskolebarn går från att lära sig läsa till att läsa för att lära. I denna fas börjar lärare förvänta sig att barnen använder sin läsförståelse för att samla kunskap om många olika ämnen. Men barn är naturligtvis inte ensamma på sin resa för att bli skickliga läsare. En hel del forsknings- och utbildningsresurser mobiliseras för att hjälpa dem på vägen. De fyra empiriska artiklarna i denna doktorsavhandling presenterar forskning ur ett kognitionsvetenskapligt perspektiv på tre aspekter som är viktiga för att förstå läsförståelse i vetenskapliga och pedagogiska sammanhang: läsaren, texten och uppgiften (Snow & RAND, 2002; van den Broek, Fletcher, & Risden, 1993).

Inom kognitionsvetenskap beskrivs en djup förståelse av en text som konstruktionen av en mental modell, en så kallad situationsmodell (Johnson-Laird, 1983; van Dijk & Kintsch, 1983). Detta innebär att läsaren måste kunna visualisera i sitt inre hur situationerna, händelserna och karaktärerna är som beskrivs i texten. Att kunna konstruera en situationsmodell beror på läsaregenskaper, textegenskaper och uppgiftskrav. För det första behövs många kognitiva färdigheter och strategier för att en läsare ska kunna konstruera en situationsmodell, exempelvis som att se samband i texten (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Graesser, Kintsch, 1994; Singer, & Trabasso, 1994). Att se samband i texten innebär att kunna läsa mellan raderna. Detta kräver att läsaren kopplar samman olika delar av texten till andra delar av texten och att delar av texten kopplas till deras tidigare kunskaper, så kallat att läsa bortom raderna. Genom att se samband berikas situationsmodellen med tidigare kunskap och läsaren förstår då att texten inte bara är en rad av ord utan att den berättar en historia med händelser som följer på varandra med orsaker och effekter. Processer och strategier, som samband, som genereras under läsning kallas online-läsprocesser. Online-läsprocesser tros bidra till läsarnas minnesrepresentation av texten. Minnesrepresentationen som uppstår då läsaren har läst klart kallas offline-textminne. Skillnader i textämne, svårighet och längd har för det andra en inverkan på vad en läsare kan förstå av texten, och påverkar därmed hur rik situationsmodellen blir som läsaren konstruerar (t.ex. van den Broek et al., 1993). Vid läsning av en enkel text, till exempel en text som använder vardagsspråk och har tydlig struktur, kan läsaren enkelt konstruera en rik situationsmodell. Därigenom uppnår läsaren en god förståelse för innehållet. Men så snart texten blir mer krävande, till exempel introducerar nya ord och begrepp, måste läsaren arbeta hårdare för att uppnå en god förståelse av texten. För det tredje kommer den uppgift som läsaren får, eller tar på sig, under läsningen att påverka förmågan att konstruera en rik situationsmodell (t.ex. van den Broek & Kendeou, 2017). Att läsa en text i sin egen takt och kunna återgå till eller på annat sätt bearbeta meningar som är svåra att förstå kommer resultera i en annan förståelse för texten än när man försöker förstå samma text men utan möjligheten att kontrollera hastigheten med vilken texten presenteras. Till exempel vid högläsning eller när du använder en digital uppläsningstjänst. Därför interagerar sannolikt de många komplexa kognitiva processerna som bidrar till läsförståelsen med läsare-, text- och uppgiftsegenskaper (Rapp & van den Broek, 2005).

Forskningen i denna avhandling fokuserar huvudsakligen på 9-12-åriga Nederländska barns förmåga att konstruera en situationsmodell av text. Detta är en viktig fas i ett barns läsutveckling. 9-12-åringar i många västerländska utbildningssystem, som i det Nederländska skolsystemet, övergång från fasen lära att läsa till fasen läsa för att lära (t.ex. Chall 1983; 1996; Poolman, Leseman, Doornenbal, & Minnaert, 2017). I denna fas, att läsa för att lära sig, förväntas barn ha blivit flytande i ordavkodningsfärdigheter och får längre texter i syfte att undervisa innehåll. Dessa texter kräver dock också ytterligare läsprocesser jämfört med de enkla texterna som användes i den första fasen. Barnens läsförståelse sjunker ofta i övergången från att lära sig att läsa till att läsa för att lära. Denna nedgång indikerar att de ökade kraven som följer med att läsa mer komplicerade texter inte är gradvis och understryker att dessa texter ställer högre krav på de unga läsarna (Chall 1983; 1996). Men med lärare och annan skolpersonal som förstår hur man läser de svårare texterna och noggrant överväger undervisningen kan barn förstå och lära även av dessa texter.

Även om skolmiljön och hemmiljön kan öka kraven på och möjligheterna för läsutveckling av barn i åldrarna 9-12, utvecklas deras kognitiva system fortfarande. Exekutiva funktioner och hjärnnätverk som stöder språkkunskaper och exekutiva funktioner fortsätter att utvecklas hos barn och långt in i tonåren (t.ex. Diamond, 2013; Gathercole, Pickering, Ambridge, Wearing, 2004; Huizinga, Dolan, & van der Molen, 2006 ; Zielinski, Gennatas, Zhou, & Seeley, 2010). Exekutiva funktioner kan beskrivas som ett paraplybegrepp för kognitiva processer som möjliggör kontroll av tankar och beteende (Diamond, 2013), till exempel arbetsminnet. Arbetsminnet kan ses som en mental arbetsyta med begränsad lagringskapacitet, som gör det möjligt att hålla informationen i tanken och manipulera denna information (Daneman & Merikle, 1996; Diamond, 2013). Arbetsminnet är viktigt för läsförståelse eftersom det ligger till grund för den kapacitet och processer som krävs för att behålla och uppdatera innehållet i berättelsen när läsaren fortsätter genom texten (t.ex. Carretti, Borella, Cornoldi och De Beni, 2009; Daneman & Merikle, 1996).

De empiriska studierna i denna avhandling ger insikt i kognitiva processer i läs- och språkförståelse hos främst barn i grundskolan. Vi undersökte hur kognitiva processer för att konstruera situationsmodeller är relaterade till läsar-, text- och uppgiftsegenskaper. Vidare beskriver vi sambandet mellan dessa tre faktorer. Lärare vet vikten av att skapa bra förutsättningar för elever. I sitt uppdrag att göra det är frågan "vad får vissa elever att lyckas och andra inte?" en vardaglig fråga. Tillsammans med tidigare forskning hjälper denna avhandling att förstå relationerna mellan de tre aspekterna för att hantera komplexiteten i undervisning och utbildningsvetenskap.

I det **andra kapitlet** studerade vi hur barn (9-11 år) skiljer sig åt i förmågan att se samband i text online och hur dessa skillnader relaterar till deras underliggande läsaregenskaper. Barn läste texter från två genrer, berättande och informativa texter. Vi identifierade tre läsprofiler som skiljer sig åt i både antal och typer av samband barnen genererar, baserat på data från när de tänkte högt medan de läste. Reflekterande Läsare genererade olika typer av samband; de kopplade samman olika textbitar med varandra, de kopplade samman texten med egen bakgrundskunskap, och förutspådde vad som skulle hända senare i texten. Beskrivande Läsare upprepade huvudsakligen texten genom att parafrasera den. De genererade även några slutsatser. Ordagranna läsare upprepade huvudsakligen texten bokstavligen; i själva verket var 60% av svaren från barn i denna profil bokstavliga textupprepningar. Ordagranna Läsare genererade få samband. På kompletterande akademiska tester visade Reflekterande Läsare betydligt högre poäng i ordläsning och icke-verbalt resonemang än Beskrivande och Ordagranna Läsare. Dessa resultat visar att både lägre kognitiva processer (som ordläsning) och högre kognitiva processer (som resonemangsförmåga) är delvis underliggande faktorer för barns skillnader i läsförståelseprocesser online. Barn visade i allmänhet samma läsprofil för både berättande och informativa texter. Generellt sett framkallade informativa texter dock färre läsprocesser där barnen visade på att se samband än vad berättande texter gjorde. Till exempel gjordes ett mindre antal förutspående samband vid läsning av informativa texter. Dessutom hittades ett större antal felaktiga samband för informativa texter än berättande texter. Eftersom att generera färre giltiga samband men fler felaktiga samband sannolikt kommer att hindra läsförståelsen, tyder dessa resultat på att barn i grundskolan fortfarande utvecklar färdigheter för att förstå informativa texter. Sammanfattningsvis är resultaten i detta kapitel i linje med tidigare forskning (t.ex. McMaster et al., 2012; Kraal et al., 2017) som har hittat läsprofiler där barn bygger antingen en situationsmodell som liknar texten eller en situationsmodell som berikas av samband. I denna studie utvidgar vi tidigare forskning genom att visa att sådana profiler finns i en grupp barn med heterogena underliggande kognitiva och språkliga förmågor, och att dessa förmågor skiljer sig mellan profilerna. Dessutom visar vi att barn producerar en mycket liknande uppsättning svar över textgenrer som resulterar i samma profiler över de olika texterna. Det fanns dock skillnader i antalet samband som genererades mellan textgenrerna. Följaktligen tyder resultaten på en viss stabilitet i barns läsförståelse med plats för situationella skillnader relaterade till textegenskaper.

I det **tredje kapitlet** studerade vi hur online-läsprocesser för barn (9-11 år) i de tre läsprofilerna (presenterades i kapitel två) relaterar till deras offline-textminne. Eftersom offline-textminnet är en viktig indikator för inlärning i skolan måste forskning om läsprofiler ta hänsyn till kvaliteten på minnesrepresentationerna. Offline-textminne av berättande och informativa texter studerades genom att undersöka om barn kom ihåg mer central information, kärnan, än perifer information från texten efter läsning, det vill säga om de visade en så kallad centralitetseffekt. Att se samband hjälper till att skapa kopplingar mellan textdelar, vilket är viktigt för att förstå kärnan, alltså huvudbudskapet, i en text. Därför förväntade vi oss att barn som genererar ett större antal samband skulle visa en större centralitetseffekt i sitt offline-textminne jämfört med barn som genererar färre samband online. Det betyder att vi antog att Reflekterande Läsare skulle visa en större centralitetseffekt än Beskrivande och Ordagranna Läsare, och att Beskrivande Läsare skulle visa en större centralitetseffekt än Ordagranna Läsare.

Barn i alla läsprofiler visade en centralitetseffekt för berättande texter. Reflekterande Läsare visade en större centralitetseffekt än Beskrivande Läsare. Varken Reflekterande Läsare eller Beskrivande Läsare skilde sig dock från Ordagranna Läsare. Vi föreslår att dessa resultat inte kan förklaras endast av antalet samband som genereras under läsning, men vi tar också hänsyn till läsarnas andra egenskaper. Reflekterande Läsare visade sig uppnå högre poäng på ordavkodning och icke-verbalt resonemang än både Beskrivande och Ordagranna Läsare. Vi föreslår att barn i varje profil engagerar sig i läsprocesserna online som passar den kognitiva kapacitet som ligger till grund för deras läsning. Även om Beskrivande Läsare får lägre poäng vid ordavkodning och icke-verbalt resonemang än Reflekterande Läsare, verkar Beskrivande Läsare försöka generera samband när de läser. Vi spekulerar i att detta kan vara en utvecklingsfas, och att försöka generera samband men samtidigt ha relativt begränsad ordläsningsförmåga och resonemangsförmåga kan möjligen hindra Beskrivande Läsare från att uppnå en offline-minnesrepresentation som liknar den som Reflekterande Läsare uppnår.

Inga centralitetseffekter hittades när barnen läste informativa texter. Sammanfattningsvis framkallade informativa texter färre samband i online-läsprocesser (andra kapitlet) och inga profilskillnader i offline-textminne (tredje kapitel). Det är möjligt att ta fram central information ur informativa texter kräver ytterligare online-läsprocesser fän bara att generera samband som våran studie inte fångade. Tidigare forskning har ofta undersökt enstaka sambandsprocesser i korta texter. Eftersom vi nu undersökte ett antal olika samband när barnen läste längre texter, fann vi att påverkan av sambandsprocesser på offline-textminne inte är tydlig, men andra läsaregenskaper i kombination med sambandsprocesser kan förklara offline-mtextinne.

I det **fjärde kapitlet** undersökte vi barns (9-12 år gamla) förmåga att använda tidsprepositionerna *innan* och *efter* när de konstruerar en situationsmodell av meningar med två händelser, exempelvis ”Innan du subtraherar ett tal, bör du lösa multiplikationen”. Sådana meningar kan vara grammatiskt komplicerade och belastar därför arbetsminnet. Tidigare forskning har beskrivit motstridiga hypoteser om arbetsminnets roll för att förstå dessa meningar. För att utöka tidigare forskning undersöker vi hur förståelse interagerar med både arbetsminneskapacitet och arbetsminnesuppdateringsförmåga. I två experiment varierade vi placeringen av *innan* och *efter* (*voordat* och *nadat* på Nederländska) i olika meningar, de stod antingen först i meningen eller i mitten. I det första experimentet bad vi deltagarna att svara på förståelsefrågan ”vad hände först?” och i det andra experimentet bad vi deltagarna att svara på förståelsefrågan ”vad hände sist?”. Genom dessa manipulationer av meningarna och uppgifterna kunde vi undersöka om förståelsen påverkades av familjariteten av prepositionen, av prepositionen position, av svarets position (huvudsats, bisats eller sista satsen) och av meningens kronologi.

Under båda experimenten fann vi att grundskolebarns läsförståelse för meningarna påverkades av huvudsatsen snarare än familjariteten av prepositionen. Barnen var känsliga för huruvida det rätta svaret på förståelsesfrågan var beläget i huvudsatsen. Särskilt det andra experimentet visade att barnens läsförståelse kvalificerades av deras arbetsminnesuppdateringsförmåga och arbetsminneskapacitet. Barn med god arbetsminnesuppdateringsförmåga visade en huvudsatsfördel, dvs. de presterade bra när rätt svar låg i huvudsatsen. Barn med låg arbetsminnesuppdateringsförmåga förstod mer när rätt svar placerades i slutet av meningen, det vill säga en ”recency-effekt”. Dessa resultat antyder kvalitativa skillnader mellan läsförståelsen för barn med låg och hög arbetsminnesuppdateringsförmåga. Dessutom påverkades barnens läsförståelse av prepositionens position i meningen och, igen, interagerade med arbetsminnets förmågor. För meningar med tidsprepositionen i början, måste läsarna hålla information om prepositionen i arbetsminnet tills de är färdiga med att läsa hela meningen och sedan tillämpa den information som prepositionen antyder. För dessa meningar ökade barnens läsförståelse ju högre kapacitet i arbetsminnet de hade. För meningar med tidsprepositionen i mitten, måste läsarna uppdatera sin mentala representation mitt i meningen. För dessa meningar ökade barnens läsförståelse ju högre uppdateringsförmåga för arbetsminnet de hade. Tillsammans indikerar dessa resultat att grundskolebarnens läsförståelse av meningar som innehåller tidsprepositionen påverkas av läsaregenskaper, där arbetsminneskapacitet och uppdatering av arbetsminne innehar olika roller.

Motstridiga hypoteser i tidigare litteratur som undersöker prepositioner härstammar från två frågor; antingen undersöktes arbetsminneskapacitet eller uppdatering av arbetsminne (resultat relaterade till läsaregenskaper) eller så undersöktes barns förståelse vid läsande respektive lyssnande av text (resultat relaterade till uppgiftsegenskaper). I kapitel fyra undersökte vi läsaregenskaper genom de två aspekterna av arbetsminnet. I det femte kapitlet undersökte vi uppgiftsegenskaper genom en läsuppgift eller en lyssningsuppgift.

I det **femte kapitlet** undersökte vi förståelse av meningar med ord med hög och låg förutsägbarhet i två metoder: läsning och lyssnande. Vi inledde denna undersökning i en vuxen befolkning (medelålder 22 år). Deltagarna läser meningar med ord med hög och låg förutsägbarhet. I följande exempel anses "akvarium" vara ett mycket förutsägbart ord "Peter tycker att tropiska fiskar är mycket vackra. Hemma har han många olika typer som simmar i ett akvarium / skål.” I följande exempel anses dock 'skål' vara ett mycket förutsägbart ord 'Pojken var mycket nöjd med sin nya guldfisk. Hemma lade han omedelbart fisken i en skål / ett akvarium med färskt vatten.” Vi undersökte huruvida uppdateringen av arbetsminnet belastas annorlunda när man läser eller lyssnar på meningar som leder till ett mycket förutsägbart eller mindre förutsägbart ord, med hjälp av ERP-komponenten N400 (N400-komponenten i ERP ger ett online mått på meningsförståelse i hjärnan medan man lyssnar), ”eye-tracking” (vid eye-tracking mäter man den tid som det tar att läsa ett ord som även det ger ett online-mått på meningsförståelse), samt ett mått för arbetsminnesuppdateringsförmåga. Vi förutspådde att en läsuppgift belastar arbetsminnet mindre än en lyssningsuppgift eftersom läsaren har möjlighet att gå tillbaka och läsa om ord som verkar mindre förutsägbara vilket inte erbjuds i samma utsträckning vid tal.

Det fanns en måttlig positiv korrelation mellan läsuppgiften och lyssningsuppgiften. Detta indikerar att förståelseprocesser i läsning och lyssnande har både gemensamma egenskaper och skillnader. För att vidare analysera arbetsminnets roll delades deltagarna in i en grupp med relativt höga poäng på arbetsminnesuppdateringsförmåga och en grupp med relativt lägre poäng på arbetsminnesuppdateringsförmåga. Resultaten indikerade att arbetsminnet är mer belastat i lyssningsuppgiften än i läsuppgiften, och förståelse är relaterad till både individuella skillnader i arbetsminnet och uppgiftskrav, såsom möjligheten att återvända till texten. I läsuppgiften såg deltagargruppen med högt arbetsminne och gruppen av deltagare med lågt arbetsminne lika länge på högt och måttligt förutsägbara ord vid första blicken. Dessutom såg båda grupperna av deltagare lika ofta tillbaka på mindre förutsägbara ord. I lyssningsuppgiften visade endast deltagare med lågt arbetsminne en uttalad N400-effekt mot mindre förutsägbara ord, vilket tyder på att de kämpar för att förstå ordet. Därav finns det en interaktion mellan läsarens egenskaper hos arbetsminnet och de uppgiftskrav som påverkar förståelsen av meningarna. Forskning om likheter och skillnader mellan att läsa och att lyssna kan ha pedagogiska fördelar eftersom uppläsningstjänster ibland används i skolor som hjälpverktyg vid läsning. Ytterligare forskning bör emellertid omfatta grundskolebarn eftersom det är svårt att göra direkta jämförelser av resultat som gäller en vuxen befolkning eftersom barns arbetsminne fortfarande utvecklas.

Sammanfattningsvis interagerar läsaregenskaper - som förmågan att se samband, ordavkodning, resonemangsförmåga och arbetsminne - med text och uppgiftsegenskaper. För att fullt ut förstå barns utveckling av läsförståelse måste forskningen beakta de kontextuella kraven, t.ex. text- och uppgiftskrav, som läsning förekommer inom. Detta innebär att utbildare måste stödja barns läsutveckling i varje nytt sammanhang.

English Summary

Reading comprehension in elementary school children:

Cognitive studies of the reader, the text, and the task

Reading comprehension is a multifaceted skillset important to acquire in order to participate in modern society; to learn at school, for work related communication, for social digitized interactions, and to keep up to date with news. Important developmental change in this skillset occurs between the ages of 9 and 12, when elementary school children go from learning to read to reading to learn. In this phase educators start expecting the children to use their reading comprehension skillset to gather knowledge about many different topics. However, children are of course not alone on their journey to become proficient readers. A great deal of research and educational resources are mobilized to help them on their way. The four empirical papers in this doctoral dissertation present research from a cognitive scientific perspective on three elements that are important to understand reading comprehension in scientific and educational contexts: the reader, the text, and the task (Snow & RAND, 2002; van den Broek, Fletcher, & Risden, 1993).

Within cognitive science, gaining deep comprehension of a text is described as the construction of a mental model, a situation model (Johnson-Laird, 1983; van Dijk & Kintsch, 1983). This means that the situations, events, and characters that are depicted in the text need to be envisioned in the reader's mind. Being able to construct a situation model depends on reader characteristics, text characteristics, and task demands. First, many cognitive skills and strategies are needed for a reader to construct a situation model, such as to make inferences from the text (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Graesser, Kintsch, 1994; Singer, & Trabasso, 1994). Inference generation entails reading between the lines. This requires the reader to connect different parts of the text to other parts of the text, and to connect parts of the text to their previous knowledge. By generating inferences, i.e. seeing how sentences are interconnected and enriching the situation model with previous knowledge, the reader understands that the text is not just a string of words, but that it tells a story with evolving events with causes and effects. Processes and strategies, such as inferences, that are generated while reading, are called online reading processes. Online reading processes are thought to contribute to the readers' memory representation of the text when having finished reading. The memory representation after having finished reading is called offline memory. Second, differences in text topic, difficulty, and length have an impact on what a reader can extract from the text, and thereby how rich the situation model becomes that they are constructing (e.g. van den Broek et al., 1993). When reading an easy text, for example a text that uses everyday language and has clear structure, the reader can easily construct a rich situation model. Thereby, the reader achieves a good understanding of the content. However, as soon as the text becomes more demanding, perhaps introducing new words and concepts, the reader needs to work harder to achieve a good understanding of the text. Third, the task that the reader is given, or takes on, while reading will have an impact on the ability to construct a rich situation model (e.g. van den Broek & Kendeou, 2017). Reading a text at own free pace and being able to revisit sentences that are hard to understand, will rule a different outcome in comprehension than when trying to understand the same text but without the ability to control the speed with which the text is presented. For example, when being read to or when using a digital read-aloud device. Hence, the many complex cognitive processes that contribute to reading comprehension, likely interact with reader, text, and task characteristics (Rapp & van den Broek, 2005).

The research in this dissertation focuses predominantly on 9-12-year-old Dutch children's abilities to construct a situation model of text. This is an important phase in a child's reading development. 9-12-year olds in many western educational systems, as in the Dutch school system, transition from the phase learning to read to the phase reading to learn (e.g. Chall 1983; 1996; Poolman, Leseman, Doornenbal, & Minnaert, 2017). In this phase, reading to learn, children are expected to have become fluent in word-decoding skills, and are given longer texts with the aim to teach content. However, these texts also require additional reading processes compared to the simple texts used in the first phase. Children's reading performance often drops in the transition from learning to read to reading to learn, a phenomenon called the fourth-grade slump. This drop indicates that the increased demands that come with reading more complicated texts is not gradual (Chall 1983; 1996). However, with educational practitioners' careful consideration of how to read the more difficult texts, children can comprehend and learn from the texts.

Although school and home environment may increase demands on and opportunities for the reading development of children in Dutch upper elementary school, their cognitive system is still developing. Executive functions, and brain networks supporting language skills and executive functions, continue to develop in older children and well into late adolescence (e.g. Diamond, 2013; Gathercole, Pickering, Ambridge, Wearing, 2004; Huizinga, Dolan, & van der Molen, 2006; Zielinski, Gennatas, Zhou, & Seeley, 2010). Executive functions can be described as an umbrella term for cognitive processes that allow control of thoughts and behavior (Diamond, 2013), such as working memory. Working memory can be seen as a mental workspace with a limited storage capacity, that enables keeping information in mind, and manipulating this information (Daneman & Merikle, 1996; Diamond, 2013). Working memory is important in the context of reading comprehension as it underlies the capacity and processes necessary to retain and update the content of the story as the reader proceeds through the text (e.g., Carretti, Borella, Cornoldi, & De Beni, 2009; Daneman & Merikle, 1996).

The empirical studies in this dissertation provide insight in cognitive processes in reading and language comprehension in primarily upper elementary school children. We examined how cognitive situation-model building processes are related to reader, text, and task characteristics. Furthermore, we describe the relations among these three factors. Educators know the importance of constructing good learning contexts for students. In their mission to do so, the question “what makes some students succeed and others not?” is an everyday issue. Together with previous research, this dissertation helps in understanding the relations between the three factors and dealing with the complexity in educational practice and educational science.

In the **second chapter** we studied how children (9-11 years old) differ in online inference generation, and how these differences relate to children’s underlying reader characteristics. Children read texts from two genres, narrative and expository texts. Based on think-aloud data, we identified three profiles of readers that differ in both the number and types of inferences they generate. Elaborating Readers generated different types of inferences; they used text connecting inferences, elaborative inferences, and predictive inferences while reading. Paraphrasing Readers predominantly repeated the text by paraphrasing it. In addition, they also generated some inferences. Literal Readers mainly repeated the text literally; in fact, 60% of responses from children in this profile were literal text repetitions. Literal Readers generated few inferences. On ancillary measures, Elaborating Readers, showed significantly higher scores on word reading and non-verbal reasoning measures than did Paraphrasing, and Literal Readers. These results show that concerning reader characteristics, both lower- order cognitive processes (such as word reading) and higher-order cognitive processes (such as reasoning ability) underpin children’s differences in online reading processes. Children generally showed the same reading profiles for both narrative and expository texts. However, generally expository texts elicited fewer elaborative reading processes than narrative texts. For example, there was a smaller number of predictive inferences made when reading expository texts. In addition, a larger number of invalid elaborative inferences were found for expository texts than narrative texts. Because generating fewer elaborative inferences and generating more invalid inferences are likely to impede comprehension, these findings indicate that upper elementary school children are still developing skills to comprehend expository texts. To sum up, the current findings are in line with previous research (e.g. McMaster et al., 2012; Kraal et al., 2017) that has found reader profiles in which children build either a situation model that closely resembles the text or a situation model that is enriched by inferences. In this study, we expand on previous research by showing that such profiles can be found in a group of children with heterogenous underlying cognitive and language abilities, and that these abilities differ between the profiles. In addition, we show that children produce a very similar set of think-aloud responses across text genres that results in the same profiles across the different texts. There were however differences in the number of inferences generated between the text genres. Hence, the results suggest a certain stability in children's text approaches with room for situational differences related to text characteristics.

In the **third chapter** we studied how online processes of children (9-11 years old) in the three reader profiles (presented in chapter two) relate to their offline text memory. Because the offline text memory is an important indicator for learning in school, research of reading profiles needs to consider the quality of memory representations. Offline memory of narrative and expository texts was studied by examining whether children remembered more central information, the gist, than peripheral information from the text after reading, i.e. whether they showed a centrality effect. Inferences help in making connections between text parts, which is important to understand the gist of a text. Therefore, we anticipated that children who generate a larger number of online inferences would show a larger centrality effect in their offline recall, compared to children who generate fewer online inferences. Meaning we hypothesized Elaborating Readers would show a larger centrality effect than Paraphrasing Readers and Literal Readers, and that Paraphrasing Readers would show a larger centrality effect than Literal Readers.

First, all groups of children showed a centrality effect for narrative texts. Elaborating Readers showed a larger centrality effect than Paraphrasing Readers. However, neither Elaborating Readers nor Paraphrasing Readers differed from Literal Readers. We suggest that these findings cannot be explained only by the number of inferences generated during reading, but we also consider reader characteristics. Elaborating readers proved to achieve higher scores on word decoding and nonverbal reasoning than both Paraphrasing, and Literal Readers. We suggest that children in each profile engage in the online reading processes that suit the cognitive capacities that underpin their reading. Although Paraphrasing Readers score lower on word decoding and reasoning tests than Elaborating Readers, Paraphrasing Readers try out some inferences while reading. Possibly, trying out inferences while having relatively underdeveloped word reading and general reasoning abilities hinders Paraphrasing Readers from achieving an offline memory representation similar to that of Elaborating Readers. Second, no centrality effects were found when the children read expository texts. In sum, expository texts elicited fewer inferences in online processes (second chapter) and no profile differences in offline comprehension (third chapter). It is possible that extracting central information in expository texts requires additional online processes to inference generation that our measures did not capture. Previous research has often examined single inference processes in short texts. As we now examined a number of different inferences when reading longer texts, we found that the affect of inference processes on offline memory is not clear-cut, but other reader characteristics combined with inference generation may explain offline memory of a text.

In the **fourth chapter** we examined children's (9-12 year old) ability to use the temporal connectives *before* and *after* when building a situation model of sentences with two events, such as "Before you subtract a number, you should solve the multiplication". Such sentences can be grammatically complex and therefore taxing for working memory. Previous research has brought contradicting hypotheses of the role of working memory for comprehending these sentences. To expand on previous research, we examine how comprehension interacts with both working memory capacity and working memory updating ability. In two experiments we varied the position of the connectives *before* and *after* (*voordat* and *nadat* in Dutch) in the sentences, in the beginning or in the middle of the sentence. In the first experiment we asked participants to answer the question "what happened first?" and in the second experiment we asked participants to answer the question "what happened last?". By these sentence and task manipulations we could investigate whether comprehension was affected by familiarity of the connective, by the position of connective, by the position of the answer (main clause, subordinate clause, or recent clause), and by sentence chronology.

Across both experiments, we found that upper-elementary school children's comprehension was affected by clause salience, rather than the familiarity of the connective. The children were sensitive to whether the correct answer to the comprehension question was situated in the main clause or the subordinate clause. Importantly, the second experiment showed that comprehension was qualified by children's working memory updating ability and working memory capacity. Children with high working memory updating ability showed a main clause advantage, i.e. they performed well when the correct answer was situated in the main clause. Children with low working memory updating ability showed a recency effect instead, i.e. their comprehension was better when the correct answer was positioned at the end of the sentence. These results suggest qualitative differences between the comprehension of children with low and high working memory updating ability. Furthermore, the position of the connective influenced comprehension and, importantly, interacted with working memory abilities. When reading sentences with sentence-initial connectives, readers need to hold information about the connective in working memory until they have finished reading the full sentence, and then apply the information that the connective implies. For these sentences, children's performance increased with a higher working memory capacity. When reading sentences with sentence-medial connectives, readers need to update their mental representation mid-sentence. For these sentences, children's performance increased with a higher working memory updating ability. Together, these findings indicate that upper-elementary school children's comprehension of sentences containing temporal connectives is affected by reader characteristics where there are dissociable contributions of working memory capacity and working memory updating.

Contradicting hypotheses in previous literature examining connectives stemmed from two issues; research either examined working memory capacity or working memory updating (results related to reader characteristics), and research either used a reading task or a listening task (results related to task characteristics). In chapter four we disentangled the relation with the two aspects of working memory. In the fifth chapter we researched task characteristics by means of a reading task or a listening task.

In the study presented in the **fifth chapter** we examined comprehension of sentences with target words of high and low predictability in two modalities: reading and listening. We started this examination in an adult population (mean age 22 years old). The participants read sentences with target words of high and low predictability. In the following example 'aquarium' is thought to be a highly predictable word "Peter thinks that tropical fishes are very beautiful. At home, he has a lot of different types swimming in an aquarium/bowl." However, in the following example 'bowl' is thought to be a highly predictable word "The boy was very happy with his new goldfish. At home he immediately put the fish in a bowl/ an aquarium with fresh water." We investigated whether working memory updating is taxed differently when reading or listening to sentences leading up to a highly predictable or less predictable word, using the ERP component N400 (The N400 component of the ERP provides an online measure of meaning processing in the brain while listening), eye-tracking (time spent on reading a word provide an online measure of meaning processing), and a working memory updating task. We predicted that a reading task taxes working memory less than a listening task as the reader has the possibility to go back and reread target words that seem less predictable.

First, there was a moderate positive correlation between the reading task and the listening task. In other words, this finding indicates that comprehension processes have both commonalities and differences in reading and listening. In addition, results indicate that working memory is more taxed in the listening task than in the reading task, and comprehension is related to both individual differences in working memory, and task demands such as the possibility to revisit the text. In the reading task, the group of participants with high working memory and the group of participants with low working memory looked equally long at highly and moderately predictable words at first gaze. In addition, both groups of participants looked back at less predictable words equally often. In the listening task, only participants with low working memory showed a pronounced N400 effect towards less predictable words, suggesting they struggle to make sense of it. Hence, there is an interaction between reader characteristics of working memory and the task demands that affect comprehension. Research concerning modality differences can have educational benefits as listening software are sometimes used in schools as reading aids. However, further research should include elementary school children because it is difficult to make direct comparisons of results concerning an adult population since children's working memory is still developing.

To sum up, reader characteristics - such as the ability to use inferences, word decoding, reasoning abilities and working memory– interact with text and task characteristics. To fully understand children's development of reading comprehension, research needs to consider the contextual demands, e.g. text and task demands, it occurs within. This implies that educational practitioners need to support children's reading development in each new context.

Reference list

Ackerman, B. P. (1986). Referential and causal coherence in the story comprehension of children and adults. *Journal of Experimental Child Psychology, 41*, 336-366.

Ackerman, B. P. (1988). Reason inferences in the story comprehension of children and adults. *Child Development, 59*, 1426-1442.

Albrecht, J. E., & O'Brien, E. J. (1991). Effects of centrality on retrieval of text-based concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17,* 5, 932-939.

Altmann, G., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33*, 583-609.

Anmarkrud, Ø., McCrudden, M. T., Bråten, I., & Strømsø H. I. (2013). Task-oriented reading of multiple documents: online comprehension processes and offline products. *Instructional Science, 41,* 5, 873-894. doi:10.1007/s11251-013-9263-8

Armbruster, B. B., Anderson, T. H., & Ostertag, J. (1987). Does text structure/summarization instruction facilitate learning from expository text? *Reading Research Quarterly, 22*, 3, 331-346.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36,* 3, 189-208. doi: 10.1016/S0021-9924(03)00019-4

Baddeley, A.D. (2007). *Working Memory, Thought and Action*. Oxford, UK: Oxford University Press.

Baddeley, A.D. (2012). Working memory: theories, models, and controversies. *The Annual Review of Psychology, 63*, 1-29.

Baker, L., & Wagner, J. L. (1987). Evaluating information for truthfulness: The effects of logical subordination. *Memory & Cognition, 15,* 3, 247-255. doi:10.3758/BF03197723

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using {lme4}. *Journal of Statistical Software, 67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology, 107*, 452-465. doi:10.1016/j.jecp.2010.05.012

Beker, K., Jolles, D., & van den Broek, P. (2017). Meaningful learning from texts: The construction of knowledge representations. In J. A. León & I. Escudero (Eds.), *Studies in Written Language and Literacy: Reading Comprehension in Educational Settings* (pp. 29-62). The Netherlands: John Benjamins Publishing Company.

Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology, 9*, 291-319. doi:10.1017/S095457949700206X

Best, R. M., Floyd, R. G., & McNamara D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology, 29,* 2, 137-164. doi:10.1080/02702710801963951

Blything, L. P., & Cain, K. (2016). Children’s processing and comprehension of complex sentences containing temporal connectives: The influence of memory on time course of accurate responses. *Developmental Psychology, 52,* 10, 1517-1529. doi:10.1037/dev0000201

Blything, L. P., Davies, R., & Cain, K. (2015). Young children's comprehension of temporal relations in complex sentences: The influence of memory on performance. *Child Development*, *86*, 1922-1934. doi: 10.1111/cdev.12412

Bogaerds-Hazenberg, S.T.M., Evers-Vermeul, J., & van den Bergh, H.H. (2020). A meta-analysis on the effects of text-structure instruction on reading comprehension in the upper elementary grades. *Reading Research Quarterly.* Advance online publication.

Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology, 39*, 3, 206-219. doi:10.1016/j.cedpsych.2014.05.003

Booth, J. R., Perfetti, C. A., & MacWhinney, B. (1999). Quick, automatic, and general activation of orthographic and phonological representations in young readers. *Developmental Psychology, 35*, 3-19.

Booth, J. R., Perfetti, C. A., MacWhinney, B., & Hunt, S. B. (2000). The association of rapid temporal perception with orthographic and phonological processing in children and adults with reading impairment. *Scientific Studies of Reading, 4*, 101-132.

Booth, J.R., Burman, D.D., Meyer, J.R., Gitelman, D.R., Parrish, T.B., Mesulam, M.M. (2002). Modality independence of word comprehension. *Human Brain Mapping, 16*, 251-61.

Boudewyn, M. A., Long, D. L., & Swaab, T. Y. (2013). Effects of working memory span on processing of lexical associations and congruence in spoken discourse. *Frontiers in Psychology, 4*. doi:10.3389/fpsyg.2013.00060

Bowyer-Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology, 75,* 189-201. doi:10.1348/000709904X22674

Britton, B. K., & Pelligrini, A. D. (1990). *Narrative Thought and Narrative Language.* Hillsdale, NJ: Earlbaum.

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior, 22*, 1-14.

Buchweitz, A., Mason, R. A., Tomitch, L., & Just, M. A. (2009). Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. *Psychology and Neuroscience*, *2*, 111-123. doi:10.3922/j.psns.2009.2.003.

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology, 103*, *2*, 429-441. doi:10.1037/a0022824

Cain, K., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal, 11*, 489-503. doi:10.1023/A:1008084120205

Cain, K., Oakhill, K., & Bryant, P. (2004) Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 91,* 1, 31-42. doi:10.1037/0022-0663.96.1.31

Calvo, M.G. (2001). Working memory and inferences: evidence from eye fixations during reading. *Memory*, *9*, 365-381.

Calvo, M.G. (2004). Relative contribution of vocabulary knowledge and working memory span to elaborative inference in reading. *Learning and Individual Differences*, *15*, 53-65.

Carlson, S. E., Seipel, B., & McMaster, K. (2014). Development of a new reading comprehension assessment: Identifying comprehension differences among readers. *Learning and Individual Differences, 32,* 40-53. doi:10.1016/j.lindif.2014.03.003

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, *19*, 2, 246-251. doi: 10.1016/j.lindif.2008.10.002

Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology, 91*, 45-66. doi:10.1016/j.jecp.2005.01.005

Clark, E. V. (1971). On the acquisition of the meaning of before and after. *Journal of Verbal Learning and Verbal Behavior, 10,* 266-275. doi:10.1016/S0022-5371(71)80054-3

Claus, B., & Kelter, S. (2006). Comprehending narratives containing flashbacks: Evidence for temporally organized representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 5, 1031-1044. doi:10.1037/0278-7393.32.5.1031

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P.G. Van Gompel, M. H. Fisher, W. S. Murray, &

R. L. Hill (Eds.) *Eye Movements: A Window on Mind and Brain* (pp. 341-372). doi:10.1016/B978-008044980-7/50017-3

Clinton, V., & van den Broek, P. (2012). Interest, inferences and learning from texts. *Learning and Individual Differences, 22,* 650-663. doi:10.1016/j.lindif.2012.07.004

Collins, L. M., & Lanza, S. T. (2010). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences.* Hoboken, NJ: Wiley.

Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., & Shankweiler, D. (2004). Sentence complexity and input modality effects in sentence comprehension: an fMRI study. *Neuroimage*, 22, 1, 11-21.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12,* 5, 769-786. doi:10.3758/BF03196772

Cook, A., Limber, J. E., & O'Brien, E. J. (2001). Situation-based context and the availability of predictive inferences. *Journal of Memory and Language, 44*, 220–234.

Cook A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes, 51,* 26 – 59. doi:10.1080/0163853X.2013.855107

Cooreman, A., & Sanford, A. (1996). Focus and syntactic subordination in discourse. Research paper no. RP-79, University of Edinburgh, HCRC.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation, *Discourse Processes, 25*, 1, 1-53. doi:10.1080/01638539809545019

Cowan, N. (1995). *Attention and Memory: An Integrated Framework*. Oxford: Oxford University Press.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review, 24,* 4, 1158-1170. doi:10.3758/s13423-016-1191-6

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*, 4, 422-433. doi. 10.3758/BF03214546

de Ruiter, L. E., Theakston, A. L., Brandt, S., & Lieven, E. V. (2018). Iconicity affects children's comprehension of complex sentences: The role of semantics, clause order, input and individual differences. *Cognition*, *171*, 202-224. doi: 10.1016/j.cognition.2017.10.015

de Leeuw, L., Segers, E., & Verhoeven, L. (2016). The effect of student-related and text-related characteristics on student's reading behavior and text comprehension: An eye movement study (2016). *Scientific Studies of Reading, 20*, 3, 248-263. doi:10.1080/10888438.2016.1146285

Debruille, J. B. (2007). The N400 potential could index a semantic inhibition. *Brain Research Reviews*, 56, 2, 472-477.

Debruille, J. B., Ramirez, D., Wolf, Y., Schaefer, A., Nguyen, T. V., Bacon, B. A., Renoult, L., & Brodeur, M. (2008). Knowledge inhibition and N400: A within-and a betweensubjects study with distractor words. *Brain Research, 1187*, 167-183.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*, 8, 1117-1121.

Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., Fletcher, J. M., & Carter, S. (2015). Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly, 50*, 4, 393-416. doi:10.1002/rrq.105

Diamond, A. (2013). Executive functions. *Annual Reviews of Psychology, 64,* 135-168. doi: 10.1146/annurev-psych-113011-143750

Diessel, H. (2005). Competing motivations for the ordering of main and adverbial clauses. *Linguistics, 43,* 3*,* 449-470. doi:10.1515/ling.2005.43.3.449.

Eason, S. H., Goldberg, L., 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,* 515-528. doi:10.1037/a0027182

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: an experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 170-189. doi: 10.1037/a0017891

Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: The current state of the literature. *Psychonomic Bulletin & Review, 22,* 1, 13-37. doi: 10.3758/s13423-014-0665-7

Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic review. *Journal of Educational Psychology, 109*, 6, 761-781. http://dx.doi.org/10.1037/edu0000180

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23.

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

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87,* 3, 215-251. doi:10.1037/0033-295X.87.3.215

Evers, G. (2008). *Programma voor berekening Cito Lees Index voor het Basisonderwijs*. P-CLIB versie 3.0. Arnhem, The Netherlands: Cito B.V.

Evers-Vermeul, J., & Sanders, T. (2009). The emergence of Dutch connectives; How cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language, 36,* 829-854. doi: 10.1017/S0305000908009227

Federmeier, K. D., & Kutas, M. (1999a). A rose by any other name: Long-term memory structure and sentence processing. *Journal of memory and Language, 41*, 469-495.

Federmeier, K. D., & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research, 8*, 373-392.

Feenstra, H., Kleintjes, F., Kamphuis, F., & Krom, R. (2010). *Begrijpend Lezen groep 3 t/m 6 LOVS*. Arnhem: Cito B.V.

Fincher-Kiefer, R. (1995). Relative inhibition following the encoding of bridging and predictive inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 981–995.

Finney, M. C., Montgomery, J. W., Gillam, R. B., & Evans, J. L. (2014). Role of working memory storage and attention focus switching in children's comprehension of spoken object relative sentences. *Child Development Research*, *2014*. doi:10.1155/2014/450734

French, L. A., & Brown, A. L. (1977). Comprehension of before and after in logical and arbitrary sequences. *Journal of Child Language, 4,* 2, 247-256. doi:10.1017/S0305000900001641

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1366-1383.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, *3*, 11, 410-419. doi: 10.1016/S1364-6613(99)01388-1

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40,* 177-190. doi: 10.1037/0012-1649.40.2.177

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, Language, Brain* (pp. 95-126). Cambridge: MIT Press.

Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences, 2,* 7, 262-268. doi:10.1016/S1364-6613(98)01187-5

Gernsbacher, M. A., & Robertson, R. R. (1999). The role of suppression in figurative language comprehension. *Journal of Pragmatics, 31*, 1619-1630.

Goldman, S. R., McCarthy, K. S., & Burkett, C. (2015). Interpretive inferences in literature. In: E. J., O'Brien, A. E., Cook, & R. F., Lorch (Eds.), *Inferences During Reading* (pp. 386-415). UK: Cambridge University Press. doi:10.1017/cbo9781107279186.018

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet, & C. E. Snow (Eds.), *Rethinking Reading Comprehension* (pp. 82-98). New York: Guilford.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48,* 1, 163-189.

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

Grunér, S., Östberg, P., & Hedenius, M. (2018). The compensatory effect of text-to-speech technology on reading comprehension and reading rate in Swedish schoolchildren with reading disability: The moderating effect of inattention and hyperactivity symptoms differs by grade groups. *Journal of Special Education Technology, 33*, 2, 98-110. doi: 10.1177/0162643417742898

Guthrie, J. T., & Alao, S. (1997). Designing contexts to increase motivations for reading. *Educational Psychologist, 32,* 95-107. doi:10.1207/s15326985ep3202_4

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied Latent Class Analysis*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511499531

Hannon, B. (2012). Understanding the relative contributions of lower-level word processes, higher-level processes, and working memory to reading comprehension performance in proficient adult readers. *Reading Research Quarterly, 47*, 2, 125-152. doi:10.1002/RRQ.013

Helder, A., van Leijenhorst, L., & van den Broek, P. (2016). Coherence monitoring by good and poor comprehenders in elementary school: comparing offline and online measures. *Learning and Individual Differences, 48*, 17-23.

Hertzum, M., Hansen, K. D., & Andersen, H. H. K. (2009) Scrutinising usability evaluation: does thinking aloud affect behaviour and mental workload?, *Behaviour & Information Technology, 28,* 2, 165-181. doi:10.1080/01449290701773842

Hickendorff, M., Edelsbrunner, P., McMullen, J., Schneider, M., & Trezise, K. (2017). Informative Tools for Characterizing Individual Differences in Learning: Latent Class, Latent Profile, and Latent Transition Analysis. *Learning and Individual Differences*, In Press. doi:10.1016/j.lindif.2017.11.001

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

Huizinga, M., Dolan, C., & van der Molen, M. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44,* 2017-2036. doi:10.1016/j.neuropsychologia.2006.01.010

Hyönä, J., Lorch Jr. R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*, 1, 44-55. doi:10.1037//0022-0663.94.1.44

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory, 18,* 4, 394-412. doi:10.1080/09658211003702171

Jobard, G., Vigneau, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2007). Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage, 34*, 2, 784-800.

Johnson-Laird, P. N. (1983). *Mental Models.* Cambridge, MA: Harvard University Press.

Jonsson, B., Wiklund-Hörnqvist, C., Nyroos, M., & Börjesson, A. (2014). Self-reported memory strategies and their relationship to immediate and delayed text recall and working memory capacity. *Educational Inquiry, 5*, 3, 385-404. doi:10.3402/edui.v5.22850

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

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Allyn & Bacon.

Kaakinen, J. K., & Hyönä, J. (2005). Perspective effects on expository text comprehension: Evidence from think-aloud protocols, eye-tracking, and recall. *Discourse Processes, 40*, 3, 239-257. doi:10.1207/s15326950dp4003_4

Karlsson, J., van den Broek, P., Helder, A., Hickendorff, M., Koornneef, A., & van Leijenhorst, L. (2018). Profiles of young readers: Evidence from thinking aloud while reading narrative and expository texts. *Learning and Individual Differences, 67,* 105–116. doi:10.1016/j.lindif.2018.08.001

Keefe, D. E., & McDaniel, M. A. (1993). The time course and durability of predictive inferences. *Journal of Memory and Language, 32*, 446–463.

Kendeou, P., Bohn-Gettler, C., White, M. J., & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading, 31*, 3, 259-272. https://doi.org/10.1111/j.1467-9817.2008.00370.x

Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice, 29*, 1, 10-16. doi:10.1111/ldrp.12025

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2007). Comprehension in preschool and early elementary children: Skill development and strategy interventions. In D. S. McNamara (Ed.), *Reading Comprehension Strategies: Theories, Interventions, and Technologies* (pp. 27-45). Mahwah, NJ: Erlbaum.

Kendeou, P., van den Broek, P. W., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*, 4, 765-778.

Kidd, E. (2013). The role of verbal working memory in children's sentence comprehension: A critical review. *Topics in Language Disorders, 33,* 208–223. doi: 10.1097/TLD.0b013e31829d623e

Kim, Y. S., & Phillips, B. (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly*, 49, 3, 269-281.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95,* 163–182. doi:10.1037/0033-295X.95.2.163

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49,* 4, 294-303. doi:10.1037/0003-066X.49.4.294

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge university press.

Klin, C. M., Guzmán, A. E., & Levine, W. H. (1999). Prevalence and persistence of predictive inferences. *Journal of Memory and Language, 40*, 593-604.

Klin, C. M., Murray, J. D., Levine, W. H., & Guzmán, A. E. (1999). Forward inferences: From activation to long-term memory. *Discourse Processes, 27*, 3, 241-260.

Knol, F. (2012). *Statusontwikkeling van wijken in Nederland 1998-2010* (Report 2012-26). Retrieved from website Sociaal en Cultureel Planbureau Den Haag: http://www.scp.nl/Publicaties/Alle_publicaties/Publicaties_2012/Statusontwikkeling_van_wijken_in_Nederland_1998_2010.

Kopatich, R. D. & Santuzzi, A. (2018, July). An Exploration of Reader Profiles by Reading Proficiency and Constructive Processes Used During Reading. Paper presentation at the 28th annual meeting of the Society for Text & Discourse.

Kraal, A., Koornneef, A. W., Saab, N., & van den Broek, P. W. (2017). Processing of expository and narrative texts by low- and high-comprehending children. *Reading and Writing,* 1-24. doi:10.1007/s11145-017-9789-2

Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). *Drie-Minuten-Toets.* Arnhem: Cito B.V.

Kucan, L., & Beck, I. L., (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of Educational Research, 67*, 271-299. doi:10.3102/00346543067003271

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203-205.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*, 463-470.

Kutas, M. & Federmeier, K.D. (2011). Thirty rears and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Landers, R.N. (2015). Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS. *The Winnower* 2. doi:10.15200/winn.143518.81744

Language and Reading Research Consortium, & Logan, J. (2016). Pressure points in reading comprehension: a quantile multiple regression analysis. *Journal of Educational Psychology.* doi:10.1037/edu0000150

Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence*, 12, 2, 131–147.

Linderholm, T. (2002). Predictive inference generation as a function of working memory capacity and causal text constraints. *Discourse Processes, 34*, 259–280.

Linderholm, T. & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology, 94*, 778-784.

Lorch, R. F. (2017). What is so difficult about expository text? In J. A. Léon & I. Escudero (Eds.), *Reading comprehension in educational settings* (pp. 145-167). Amsterdam, The Netherlands: John Benjamins Publishing Company. doi:10.1075/swll.16.06lor

Lynch, J. S., van den Broek, P., Kremer, K. E., Kendeou, P., White, M. J., & Lorch, E. P. (2008). The development of narrative comprehension and its relation to other early reading skills. *Reading Psychology*, *29*, 4, 327-365. doi:10.1080/02702710802165416

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21,* 3, 251-283. doi: 10.1207/S1532690XCI2103_02

Magnusson Amu, N. (2020). Att lyssna sig till kunskap. *Digitalare undervisning, 1,* 18-19. https://www.nok.se/contentassets/e84656834f6d4a5aa4a556f4093b8ba9/digital_undervisning_1_2020_webb.pdf

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9,* 111-151.

Mann, W. C., & Thompson, S. A. (1986). Relational propositions in discourse. *Discourse Processes, 9,* 1, 57-90. doi:10.1080/01638538609544632

Mayer, R. E. (1996). Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. *Educational Psychology Review, 8*, 4, 357-371. doi:10.1007/BF01463939

McCrudden, M. T., & Kendeou, P. (2014). Exploring the link between cognitive processes and learning from refutational text. *Journal of Research in Reading, 37*, S1, S116-S140. doi:10. 1111/j.1467-9817.2011.01527.x

McGee, A., & Johnson, H. (2003). The effect of inference training on skilled and less skilled comprehenders. *Educational Psychology, 23*, 1, 49-59. doi:10.1080/01443410303220

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 12*, 82-91.

McMaster, K. L., Van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., Bohn-Gettler, C. M., & Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences, 22,* 100-111. doi:10.1016/j.lindif.2011.11.017

McNamara, D. S. (2004) SERT: Self-Explanation Reading Training. *Discourse Processes, 38,* 1, 1-30, doi: 10.1207/s15326950dp3801_1

McNamara, D. S., Boonthum, C., Kurby, C. A., Magliano, J., Pillarisetti, S., & Bellissens, C. (2009). Interactive paraphrase training: The development and testing of an iSTART module. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Artificial Intelligence in Education, Building Learning Systems That Care: from Knowledge Representation to Affective Modelling* (pp. 181-188). The Netherlands, IOS Press.

McNamara, D. S., Floyd, R. G., Best, R., & Louwerse M. M. (2004). World knowledge driving young readers' comprehension difficulties. *Proceedings of the 6th International Conference on Learning Sciences,* 326-333

McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes, 22,* 247-287. doi:10.1080/01638539609544975

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. Ross (Ed.) *The Psychology of Learning and Motivation* (pp. 297-384). Burlington: Academic Press.

McNamara, D. S., Ozuru, Y. & Floyd, G. R. (2011). Comprehension challenges in the fourth grade: The roles of text cohesion, text genre, and readers' prior knowledge. *International Electronic Journal of Elementary Education, 4*, 1, 229-257.

Medina, A. L., & Pilonieta, P. (2006). Once upon a Time: Comprehending Narrative Text. In J. S. Schumm (Ed.), *Reading Assessment and Instruction for all Learners* (pp. 222-261). New York: Guilford Press.

Meyer, B. J. F. (1985). Prose analysis: Purposes, procedures, and problems. In B. K. Britten & J. B. Black (Eds.), *Understanding Expository Text: A Theoretical and Practical Handbook for Analyzing Explanatory Text* (pp. 11-64). Hillsdale, NJ: Erlbaum.

Meyer, B. J. F. (1987) Following the author's top-level organization: An important skill for reading comprehension. In R. J. Tierney, P. L. Anders, & J. N. Mitchell (Eds.), *Understanding Readers' Understanding* (pp. 59-76). Hillsdale, NJ: Lawrence Erlbaum Associates.

Meyer, B. J. F., Brandt, D.M., & Bluth, G.J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth grade students. *Reading Research Quarterly, 16*, 1, 72-103. doi:10.2307/747349

Meyer, B. J., & Ray, M. (2011). Structure strategy interventions: Increasing reading comprehension of expository text. *International Electronic Journal of Elementary Education, 4*, 127-152.

Michael, E. B., Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). fMRI investigation of sentence comprehension by eye and by ear: Modality fingerprints on cognitive processes. *Human Brain Mapping, 13*, 239-252.

Miller, A. C., & Keenan, J. M. (2009). How word reading skill impacts text memory: The centrality deficit and how domain knowledge can compensate. *Annals of Dyslexia*, *59*, 2, 99–113. doi:10.1007/s11881-009-0025-x

Miller, A. C., & Keenan, J. M. (2011). Understanding the centrality deficit: Insight from foreign language learners. *Memory & Cognition*, *39,* 5, 873-883. doi:10.3758/s13421-010-0062-z

Miltsakaki, E. (2003). *The Syntax-Discourse Interface: Effects of the Main-Subordinate Distinction on Attention Structure* (Doctoral dissertation). Retrieved from Repository University of Pennsylvania. (AAI3087438)

Miyake, A., & Shah, P. (Eds.). (1999). *Models of Working Memory; Mechanisms of Active Maintenance and Executive Control.* New York, NY: Cambridge University Press.

Mo, L., Chen, H., Li, Y., Chen, Z., & He, X. (2007). Effects of event-related centrality on concept accessibility. *Discourse Processes, 43,* 3, 229-254. doi:10.1080/01638530701226204

Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Letters to Nature, 395,* 71-73. doi:10.1038/25731

Naglieri, J. A. (2001). Do ability and reading achievement correlate? *Journal of Learning Disabilities, 34*, 4, 304-305.

Narvaez, D., van den Broek, P., & Barrón Ruiz, A., (1999). Reading purpose, type of text and their influence on think-alouds and comprehension measures. *Journal of Educational Psychology, 91,* 3, 488-496. doi:10.1037/0022-0663.91.3.488

Nation, K., & Snowling, M. J. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67*, 359-370. doi:10.1111/j.2044-8279.1997.tb01250.x

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39,* 85-101. doi:10.1006/jmla.1998.2564

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70,* B1-B13. doi:10.1016/S0010-0277(99)00004-9

Natsopoulos, D., & Abadzi, H. (1986). Understanding linguistic time sequence and simultaneity: A literature review and some new data. *Journal of Psycholinguistic Research, 15,* 3, 243-273.

Naveh-Benjamin, M., Moscovitch, M., & Roediger, H. L. (2001). *Perspectives on Human Memory and Cognitive Aging: Essays in Memory of Fergus Craik.* New York, NY: Psychology Press.

Nieuwland, M. S. (2015). The truth before and after: brain potentials reveal automatic activation of event knowledge during sentence comprehension. *Journal of Cognitive Neuroscience, 27*, 2215-2228.

Nieuwland, M.S. (2016). Quantification, prediction and the online impact of sentence truthvalue: evidence from event-related potential. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 316-334.

Noordman, L., Vonk, W., & Kempf, H. (1992). Causal inferences during reading of expository texts. *Journal of Memory and Language, 31,* 573-590. doi:10.1016/0749-596X(92)90029-W

O'Brien, E. J., Cook, A. E., & Lorch, Jr. R. F. (2015). *Inferences during reading*. UK: Cambridge University Press. doi:10.1017/CBO9781107279186

O'Brien, E. J., & Myers, J. L. (1987). The role of causal connections in the retrieval of text. *Memory & Cognition, 15,* 419–427.

Oakes, J. (2008). Keeping track: Structuring equality and inequality in an era of accountability. *Teachers College Record, 110*, 3, 700-712.

Oakhill, J., Cain, K., & Elbro, C. (2015). *Understanding and Teaching Reading Comprehension: A Handbook.* UK: Routledge.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence,31*, 167-193.

Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In J. Robertson, & M. Kaptein (Eds.), *Modern Statistical Methods for HCI* (pp. 275-287). Springer International Publishing. doi:10.1007/978-3-319-26633-6_12

Olson, M. W. (1985). Text type and reader ability: The effects on paraphrase and text-based inference questions. *Journal of Reading Behavior, 3*, 199-214. doi:10.1080/10862968509547540

Omanson, R. C. (1982a). An analysis of narratives: Identifying central, supportive, and distracting content. *Discourse Processes, 5,* 195–224.

Omanson, R. C. (1982b). The relation between centrality and story category variation. *Journal of Verbal Learning and Verbal Behavior, 21*, 3, 326–337.

Otten, M., Nieuwland, M. S., & Van Berkum, J. J. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC neuroscience, 8,* 1, 89.

Otten, M., & van Berkum, J.J.A. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research, 1291*, 92-101.

Pander Maat, H., Kraf, R., & Dekker, N. (2017). *Handleiding T-Scan.* Retrieved from https://webservices-lst.science.ru.nl/tscan/ on July 17th, 2018.

Pickering, M.J., & Traxler, M.J. (1998). Plausibility and recovery from garden paths: an eyetrackingstudy. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 7278-7393.

Poolman, B. G., Leseman, P. P. M., Doornenbal, J. M., & Minnaert, A. E. M. G. (2017). Development of the language proficiency of five- to seven-year-olds in rural areas. *Early Child Development and Care, 187,* 3-4, 756-777. doi:10.1080/03004430.2016.1203787

Pressley, M., & Afflerbach, P. (1995). *Verbal Protocols of Reading: The Nature of Constructively Responsive Reading.* Hillsdale, NJ: Erlbaum.

Pyykkönen, P., & Järvikivi, J. (2012). Children and situation models of multiple events. *Developmental Psychology, 48, 2,* 521-529. doi:10.1037/a0025526

R Core Team. (2017). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org/

Rapp, D. N., & Mensink, M. C. (2011). Focusing effects from online and offline reading tasks. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text Relevance and Learning from Text* (pp. 141-164). Charlotte, NC, US: IAP Information Age Publishing.

Rapp, D. N., & van den Broek, P. (2005). Dynamic text comprehension: An integrative view of reading. *Current Directions in Psychological Science, 14,* 5, 276-279. doi:10.1111/j.0963-7214.2005.00380.x

Rapp, D. N., van den Broek, P., McMaster, K.L., Kendeou, P., & Espin, C.A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading, 11,* 289-312. doi:10.1080/10888430701530417

Raven, J. (2008). General introduction and overview: The Raven Progressive Matrices Tests: Their theoretical basis and measurement model. In J. Raven & J. Raven (Eds.) *Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics*. (pp. 17-68). Unionville, New York: Royal Fireworks Press.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview.* San Antonio: Harcourt Assessment.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422

Rayner, K., & Sereno, S. C. (1994). Eye movements in reading: Psycholinguistic studies. In Gernsbacher, Morton Ann (Ed), (1994). *Handbook of Psycholinguistics.* (pp. 57-81). San Diego, CA, US: Academic Press.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S.P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Leaning, Memory, and Cognition, 30*, 1290-1301.

Samuelstuen, M. S., & Bråten, I. (2005). Decoding, knowledge, and strategies in comprehension of expository text. *Scandinavian Journal of Psychology, 46,* 107-117. doi:10.1111/j.1467-9450.2005.00441.x

Sanford, A. J. (2002). Context, attention and depth of processing during interpretation. *Mind and Language, 17,* 188-206. doi:10.1111/1468-0017.00195

Schellings, G., Aarnoutse, C., & van Leeuwe, J. (2006). Third-grader's think-aloud protocols: Types of reading activities in reading an expository text. *Learning and Instruction, 16,* 6, 549-568. doi:10.1016/j.learninstruc.2006.10.004

Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL.* Amsterdam: Harcourt Test Publishers.

Schmiedek, F., Hildebrandt, A., Lövdén, M., Lindenberger, U., & Wilhelm, O. (2009). Complex span versus updating tasks of working memory: the gap is not that deep. Journal of Experimental Psychology. *Learning, Memory, and Cognition, 35,* 4, 1089-1096. doi:10.1037/a0015730.

Schneider, W., Kron-Sperl, V., & Hünnerkopf, M. (2009). The development of young children’s memory strategies: Evidence from the Würzburg Longitudinal Memory Study. *European Journal of Developmental Psychology, 6*, 1, 70-99. doi: 10.1080/17405620701336802.

Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology:Learning, Memory, and Cognition, 14*, 344–354.

Seipel, B., Carlson, S. E., & Clinton, V. E. (2017). When do comprehender groups differ? A moment-by-moment analysis of think-aloud protocols of good and poor comprehenders. *Reading Psychology, 38,* 1-32. doi:10.1080/02702711.2016.1216489

Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition, 20,* 5, 539-548. doi:10.3758/BF03199586

Singer, M., & Ritchot, K. F. M. (1996). The role of working memory capacity and knowledge access in text inference processing. *Memory & Cognition, 24*, 733–743.

Snow, C. E. & RAND (2002). *Reading for Understanding: Toward an R&D Program in Reading Comprehension.* Rand Corporation, CA, USA.

Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: evidence from noun-noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 1162-1169.

St. George, M., Mannes, S., & Hoffman, J. E. (1997). Individual differences in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience, 9*, 776–787.

Sturt, P. (2007). Semantic re-interpretation and garden path recovery. *Cognition, 105*, 477-488.

Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). *Language-related ERP components. Oxford handbook of event-related potential components*, 397-440.

Swanson, H. L. (1992). Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology, 84*, 4, 473.

Swanson, H. L., Cochran, K. F., & Ewers, C. A. (1989). Working memory in skilled and less skilled readers. *Journal of Abnormal Child Psychology, 17,* 2, 145-156.

Säljö, Roger (2015). Lärande. En introduktion till perspektiv och metaforer. Sweden: Gleerups.

Taylor, B. (1980). Children's memory for expository text after reading. *Reading Research Quarterly, 15*, 399-411.

Taylor, B. (1982). Text structure and children's comprehension and memory for expository materials. *Journal of Educational Psychology, 15*, 401-405.

Tellings, A., Hulsbosch, M., Vermeer, A., & van den Bosch, A. (2015). BasiLex: an 11.5-million words corpus of Dutch texts written for children. *Computational Linguistics in the Netherlands Journal, 4,* 191-208.

Till, R. E., Mross, E. F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory & Cognition, 16*, 283-298.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes, 21*, 3, 255-287. doi.org/10.1080/01638539609544959

Trosborg, A. (1982). Children's comprehension of before and after reinvestigated. *Journal of Child Language, 9*, 2, 381-402. doi:10.1017/S0305000900004773

Tun, P. A. (1989). Age differences in processing expository and narrative text. *Journal of Gerontology: Psychological Sciences, 44*, 9–15.

Turley-Ames, K. & Whitfield, M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language, 49,* 4, 446-468. doi: 10.1016/S0749-596X(03)00095-0.

van Berkum, J. J., Hagoort, P., & Brown, C. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*, 657-671.

van den Broek, P. W. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In: D. A. Balota, G. B. Flores d'Arcais & K. Rayner (Eds.), *Comprehension Processes in Reading* (pp. 423-446). Hillsdale, NY: Lawrence Erlbaum Associates.

van den Broek, P. W. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science, 328*, 453-456.

van den Broek, P., Espin, C., McMaster, K., & Helder, A. (2017). Developing reading comprehension interventions – Perspectives from theory and practice. In E. Segers & P. van den Broek (Eds.), *Developmental Perspectives in Written Language and Literacy* (pp. 85-101). Amsterdam, Philadelphia: John Benjamins Publishing Company. Doi:10.1075/z.206

van den Broek, P., Fletcher, C. R., & Risden, K. (1993). Investigations of inferential processes in reading: A theoretical and methodological integration. *Discourse Processes*, 16, 169–180.

van den Broek, P., Fletcher, C., & Risden, K. (1993). Investigations of inferential processes in reading: A theoretical and methodological integration. *Discourse Processes, 16,* 169-180.

van den Broek, P. W., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology, 22,* 3, 335-351. doi:10.1002/acp.1418

van den Broek, P., & Kendeou, P. (2017). Development of reading comprehension: Change and continuity in the ability to construct coherent representations. In K. Cain, D.L. Compton, & R. K. Parrila (Eds.), *Theories of Reading Development* (pp. 283-305). Amsterdam, The Netherlands: John Benjamins Publishing Company.

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

van den Broek, P.W., Risden, K., Fletcher, C. R., & Thurlow, R. (1996) A 'landscape' view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In: Britton, B.K. & Graesser, A.C. (Eds.), *Models of Understanding Text* (pp. 165-187). Hillsdale, NJ: Erlbaum.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The Construction of Mental Representations During Reading* (pp. 62-87). Mahwah, New Jersey: Lawrence Erlbaum Associates.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension.* New York: Academic Press.

Van Petten, C. V. (1993). A comparison of lexical and sentence-level context effects in event related potentials. *Language and Cognitive Processes, 8*, 485-531.

van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes, 52,* 1, 47-76. doi:10.1080/0163853X.2014.905237

Vermunt, J. K., & Magidson, J. (2013). *Technical Guide for Latent GOLD 5.0: Basic, Advanced, and Syntax.* Belmont, MA: Statistical Innovations Inc.

Vermunt, J. K., Tran, B., & Magidson, J. (2008). Latent class models in longitudinal research. In S. Menard (Eds.), *Handbook of Longitudinal Research: Design, Measurement, and Analysis* (pp. 373-385). Burlington, MA: Elsevier.

Virtue, S., van den Broek, P., & Linderholm, T. (2006). Hemispheric processing of inferences: The effects of textual constraint and working memory capacity. *Memory & Cognition, 34*, 6, 1341-1354. doi: 10.3758/BF03193276

Warren, T., McConnell, K., Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology:Leaning, Memory, and Cognition, 34*, 1001-1010.

Weekers, A., Groenen, I., Kleintjes, F., & Feenstra, H. (2011). *Begrijpend Lezen groep 7 en 8 LOVS.* Arnhem: Cito B.V.

Weingartner, K. M., Guzmán, A. E., Levine, W. H., & Klin, C. M. (2003). When throwing a vase has multiple consequences: Minimal encoding of predictive inferences. *Discourse Processes, 36*, 131–146.

Whitney, P., Ritchie, B.G., & Clark, M.B. (1991). Working memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes, 14*, 133–145.

Whitney, P., Ritchie, B. G., & Crane, R. S. (1992). The effect of foregrounding on readers' use of predictive inferences. *Memory & Cognition, 20*, 4, 424-432.

Williams, J. P., Hall, K. M., & Lauer, K. D. (2004). Teaching expository text structure to young at-risk learners: Building the basics of comprehension instructions. *Exceptionality, 12*, 3, 129-144. doi:10.1207/s15327035ex1203_2

Williams, J. P., & Pao, L. S., (2011). Teaching narrative and expository text structure to improve comprehension. In R. E. O'Connor & P. F. Vadasy (Eds.), *Handbook of Reading Interventions* (pp. 254-278). New York, NY: The Guilford Press.

Willingham, D. T. (2015). *Raising Kids Who Read: What Parents and Teachers can Do.* San Francisco, CA: Jossey-Bass.

Wolfe, M. (2005). Memory for narrative and expository text: Independent influences of semantic associations and text organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 359–364. doi:10.1037/0278-7393.31.2.359

Wolfe, M. B., & Woodwyk, J. M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology*, *80*, 3, 341-362. doi:10.1348/000709910X485700

Wood, S. G., Moxley, J. H., Tighe, E. L., & Wagner, R. K. (2018). Does use of text-to-speech and related read-aloud tools improve reading comprehension for students with reading disabilities? A meta-analysis. *Journal of Learning Disabilities, 51,* 1, 73-84. doi: 10.1177/0022219416688170.

Ye, Z., Kutas, M., St George, M., Sereno, M. I., Ling, F., & Münte, T. F. (2012a). Rearranging the world: Neural network supporting the processing of temporal connectives. *NeuroImage, 59,* 4, 3662-3667. doi:10.1016/j.neuroimage.2011.11.039

Ye, Z., Milenkova, M., Mohammadi, B., Kollewe, K., Schrader, C., Dengler, R., Samii, A., & Münte, T. F. (2012b). Impaired comprehension of temporal connectives in Parkinson's disease - A neuroimaging study. *Neuropsychologia, 50,* 8, 1794-1800. doi:10.1016/j.neuropsychologia.2012.04.004

Yeari, M., Oudega, M., & van den Broek, P. (2016). The effect of highlighting on processing and memory of central and peripheral text information: evidence from eye movements. *Journal of Research in Reading,* 1-19. doi:10.1111/1467-9817.12072

Yeari, M., van den Broek, P., & Oudega, M. (2015). Processing and memory of central versus peripheral information as a function of reading goals: Evidence from eye-movements. *Reading and Writing, 28*, 8, 1071–1097. doi:10.1007/s11145-015-9561-4

Yekovich, F. R., & Walker, C. H. (1986). Retrieval of scripted concepts. *Journal of Memory and Language*, *25,* 5, 627-644.

Zabrucky, K., & Ratner, H. H. (1992). Effects of passage type on comprehension monitoring and recall in good and poor readers. *Journal of Literacy Research*, *24,* 3, 373-391.

Zielinski, B. A., Gennatas, E. D., Zhou, J., & Seeley, W.W. (2010). Network-level structural covariance in the developing brain. *PNAS, 107*, 42, 18191-18196. doi:10.1073/pnas.1003109107

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22,* 5, 1196-1207.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology of Learning, Memory, and Cognition, 21,* 386-397.

Propositions

- Children’s reading behavior can be characterized as belonging to different reading profiles based on the number and types of inferences they generate (this dissertation)
- The use of inferences during reading is partially dependent on children’s underlying reading and cognitive abilities (this dissertation)
- The characterization of reading profiles is stable across text genres but not across think-aloud and recall tasks (this dissertation)
- Working memory capacity and working memory updating make separate contributions to children’s reading comprehension (this dissertation)
- Examining both working memory capacity and updating in children helps to understand how reading models apply to developing readers.
- When you are not satisfied with the answer, rephrase your question to emphasize the ‘re’ in research.
- Reading and listening to texts make partially different demands on working memory which impacts comprehension differently (this dissertation)
- Readers in different profiles, based on various data, differ in whether they focus predominantly on local coherence or try to comprehend global coherence of text.
- Educational sciences and practices need to account for interactions of reader, text, and task to provide good explanations of reading development.
- Educators in schools need to be given the means to get knowledgeable about new research to find the solutions necessary in their daily work and function as a sounding board to the reasearch community.

Dank aan / Thanks to:

Promotor en co-promotor Paul en Linda, voor het vertrouwen en de vrijheid die jullie mij hebben gegeven, voor de inspiratie en scherpe blik op het onderzoek. De kans die jullie mij hebben gegeven is van onschatbaar waarde.

De commissie, Prof. Denessen, Dr. Sikkema-de Jong, Prof. Segers, Dr. Evers-Vermeulen, Prof. Rispens, Dr. Veerbeek, hartelijk dank voor het lezen en evalueren van dit proefschrift.

Alle scholen, leerkrachten en leerlingen die het onderzoek in deze dissertatie mogelijk hebben gemaakt door een warm welkom en veel interesse in het onderzoek te tonen.

Claudia Roos, Erna den Besten, Rosalie Botman, Willeke Buijs, Ghislaine le Large, and Tanja Verhaar voor hun hulp bij het verzamelen van de data.

De paranimfen, Marja en Amy, dank voor alles. Voor ideeën over voeten in een bak koud water stoppen tijdens hete dagen op FSW, koffies regelen, en een klankboard zijn zowel op inhoudelijk vlak als tijdens het proces.

Wetenschappelijke zussen, Anne dank voor het mij letterlijk en soms figuurlijk uit het water helpen! Astrid, Jolien, Karin, Katinka, Marcella, en Marloes voor alle dansjes, reizen, humor en vragen naar de onderzoeksvraag! Wat een ontzettend fijne steun.

Leiden collega's, Arnout, Dietsje, Folke, Kim, Lesya, Maartje, Marian, voor altijd paraat staan met hulp en advies, uitdaging en steun. Anita, Marleen, Terry, Marleen, en Alissa, voor alle gezelligheid en het dragelijk maken van de zware momenten. Ook een bijzonder dank aan Chris en haar AIOs.

Mamma och pappa för att ni alltid finns där, no matter how far away I've been you have always been ready to jump in. It gives true peace of mind to know I can always fall back on you.

Brorsan, min förebild, Anna, Felix, och Mons, ni är ljusen i mitt liv, så mycket roligt vi delar!

Farmor och hela tjocka släkten, så mycket kärlek, vilken glädje!

Anna and Colin, thank you for designing the cover, and for all the positivity and fun even in hard times.

Sisters from other misters, Lina, Nancy, Ida, Jenny, thanks for sharing EVERYTHING in life! Massa kärlek till era fina familjer!

Kickboxdames, Annemarieke, Amy, Lisa, Lysanne, Sigrid, met zo veel Brains and Biceps hebben we enorm veel lol samen!

De Hickendorff- familie, Annelies, Marian, en de kids wat een fijne connectie tussen het leven in Nederland en Zweden. Dank voor al jullie steun en gezelligheid!

My international family away from home, Andreea, Erdinc, Ricardo, Archontia, Sofia, Maria, Paula, Annemarie for keeping me updated, my eyes on the horizon, letting me feel grounded where ever I am, for all the love and laughters we share.

Särskilt tack till ALLA kollegor vid Örebro universitet, att man kan få ha det så bra!

Michel voor de aanzet om naar Nederland te gaan, dank ook aan je familie en vrienden.

Adham for cheering me on when crossing the finishing line, for your warmth and laugther.

CURRICULUM VITAE

Josefine Karlsson was born on August 2nd, 1984, in Ryd, Skövde, Sweden. After graduating from secondary school at Västerhöjdsgymnasiet (Skövde), Josefine obtained her Bachelor´s degree in Cognitive Sciences in 2007 at Skövde Högskola, and her Research Master´s degree in Cognitive Sciences in 2011 at Universiteit van Amsterdam. During her research master program, she developed an interest in research of reading and language development.

In March 2012 Josefine started her PhD project at the Department of Educational Studies at the Institute of Education and Child Studies. Under the supervision of Prof. Dr. Paul van den Broek and Dr. Linda van Leijenhorst she worked on several research projects, including those reported in this dissertation. In addition, Josefine was involved in various teaching activities for students at different programs and levels and thereby obtained her University Teaching Qualification (Basiskwalificatie Onderwijs: BKO).

In 2018 Josefine joined Örebro Universitet as a lecturer in pedagogy at School of Humanities, Education and Social Sciences. In Örebro Josefine has joined the research environment Special Education, Development, and Learning. In addition, she has received three grants to develop her social innovation, the Digital Reading Companion, and promote and do research regards digitalization of schools and education within the platforms Social Impact Lab and Teacher Education of Tomorrow.

List of publications

Karlsson, J., Jolles, D., Koornneef, A., van den Broek, P., & van Leijenhorst, L. (2019). Individual differences in children's comprehension of temporal relations: Dissociable contributions of working memory capacity and working memory. *Journal of Experimental Child Psychology, 185,* 1-18.

Karlsson, J., van den Broek, P., Helder, A., Hickendorff, M., Koornneef, A., & van Leijenhorst, L. (2018). Profiles of young readers: Evidence from thinking aloud while reading narrative and expository texts. *Learning and Individual Differences, 67,* 105-116.

Helder, A., van den Broek, P., Karlsson, J., & van Leijenhorst, L. (2017). Neural correlates of coherence-break detection during reading of narratives. *Scientific Studies of Reading, 21, 6,* 463-479.

Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice, 29*, 10-16.

Reading comprehension is a multifaceted skillset important to acquire in order to participate in modern society; to learn at school, for work related communication, for social digitized interactions, and to keep up to date with news. Important developmental change in this skillset occurs between the ages of 9 and 12, when elementary school children go from learning to read to reading to learn. In this phase, children are expected to start using their reading comprehension skillset to gather knowledge about many different topics. However, children are of course not alone on their journey to become proficient readers. A great deal of research and educational resources are mobilized to help them on their way. With this doctoral dissertation I aim to contribute to the scientific knowledge of reading comprehension and educational practices. This gathering of four empirical papers presents research from a cognitive scientific perspective on three elements that are important to understand reading comprehension in scientific and educational contexts: the reader, the text, and the task.

Josefine Karlsson
Photo: Jesper Eriksson, Örebro universitet