How Reader and Task Characteristics Influence Young Readers' Comprehension Monitoring

Sabrina Ammi

Thesis submitted for the Degree of Doctor of Philosophy

Lancaster University

September 2015

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank the schools and children who participated in this research. Particular thanks to the members of staff who were so welcoming and the individual children who brightened my days.

I would also like to thank Prof. Kate Cain for her supervision and feedback on drafts of this work. Thanks also to Barrie Usherwood for his technical support (and patience!).

My PhD was funded by a Faculty of Science and Technology studentship. I am thankful for the financial support provided by this award.

I would also like to thank the academic community for their support and encouragement. Particular thanks to the special few for the banter and laughs.

Special thanks to my Mom, Papa and Sister for taking this journey with me. I am very grateful for their continual love, belief and support.

Declaration

I declare that this thesis is my own work and has not been previously submitted for any other degree or qualification.

Abstract

Comprehension monitoring is defined as the process by which reader's evaluate and regulate their understanding of text (e.g., Baker, 1985). Comprehension monitoring is an important component skill of reading comprehension (e.g., Cain, Oakhill & Bryant, 2004). Despite the importance of comprehension monitoring in reading comprehension, relatively little research has been undertaken to explore the development of comprehension monitoring or the task and reader characteristics critical to the development of this skill. To address this gap in the literature, this thesis explores the development of comprehension monitoring in children aged 7 to 10 years. A series of experiments are presented which explore monitoring of nonwords, general knowledge violations and internal inconsistencies using off-line and real-time measures. Experiments also explore the relationship between monitoring and working memory capacity.

Findings reveal developmental differences in comprehension monitoring. Older children are better at correctly judging the sense of information and more likely to adjust their reading behaviour in relation to error information. It seems that both age groups undertake similar monitoring behaviours, albeit with different levels of success. A range of task and reader characteristics influence monitoring skill. Findings demonstrate that task instructions influence reading behaviour. Children undertake a more purposeful and careful reading of the text when alerted that texts may contain errors. Findings also demonstrate differences in children's proficiency in adopting standards of evaluation. Children encounter most difficulties in adopting the internal consistency standard, perhaps because this standard requires children to integrate and compare the comprehensibility of information at the text-level. In addition, within error manipulations demonstrate that children use the explicitness of error information as a criterion for monitoring comprehension. Further, findings reveal that the relationship between comprehension monitoring and working memory capacity is relatively weak. Interestingly, these findings question the importance of working memory capacity as a source of monitoring difficulties. In the context of the situation model, these findings suggest that monitoring difficulties may arise from failures in constructing a richly elaborated situation model, rather than failures in updating the situation model.

CHAPTER ONE

THESIS SUMMARY

This thesis explores the development of comprehension monitoring in children aged 7 to 10 years, and the critical task and reader characteristics that influence this skill. Comprehension monitoring is defined as the process by which reader's evaluate and regulate their understanding of text (e.g., Baker, 1985). Research has identified that monitoring is an important component process of reading comprehension, this may be due to monitoring processes featuring in the construction of the situation model (e.g., Cain, Oakhill & Bryant, 2004; Rubman & Waters, 2000). To construct a coherent and richly connected situation model, dimensions of the text must be monitored and relevant information integrated and updated into the evolving model. Despite the importance of comprehension monitoring in reading comprehension, relatively little research has been undertaken to explore the development of comprehension monitoring or the task and reader characteristics critical to the development of this skill. To address this gap in the literature, this thesis presents a series of experiments that explore children's comprehension monitoring during the period of transition between beginning and fluent reading. Experiments explore different aspects of comprehension monitoring including the monitoring of different error types, monitoring in response to instructions that do and do not alert to the task of evaluating the sense of information, and the relationship between monitoring and independent skills such as working memory capacity. To account for the multi-dimensional nature of comprehension monitoring, this thesis utilises both off-line and real-time measures.

Chapter 2 provides a review of the literature as a background to the experimental work conducted. First, an overview of the relationship between comprehension monitoring in reading comprehension is provided. This includes discussion of the role of monitoring in the construction of the situation model, the nature of the monitoring process, and the role of monitoring in current models and theories of reading comprehension. Second, characteristics of the task that may influence comprehension monitoring are discussed. These characteristics include the type of error a participant must detect and the nature of instructions given to participants that inform them of the need to evaluate the sense of information. Third, characteristics of the reader that may influence comprehension monitoring are discussed. These characteristics include children's knowledge and use of different reading goals, standards of evaluation and working memory capacity. Finally, ways to assess comprehension monitoring are discussed. The advantages and disadvantages of both off-line (e.g., underlining of text and verbal responses to questions) and real-time (e.g., self-paced reading times and eye-movements) measures are discussed.

Chapter 3 explores children's off-line monitoring of nonword (i.e., a group of letters that look like a word but are not accepted as such by native speakers of a language), general knowledge violation (i.e., information that violates common world knowledge) and internal inconsistency (i.e., information in one sentence contradicts information presented in a previous sentence) errors. Children were required to underline errors in texts and respond to the question 'did the story make sense?' by ticking a sad or happy cartoon face. Developmental differences were found: older children correctly identified more errors and made more correct sense judgements than younger children. Interestingly, younger children underlined more text than older children. This suggests that younger children's poorer monitoring does not result from a reluctance to point out and underline problems within texts. Rather, younger children attempt to undertake the task but their monitoring is less accurate than that of older children. Also, monitoring differences across the error types were found: children were most accurate at monitoring nonwords and least accurate at monitoring internal inconsistencies. This suggests that children are able to proficiently adopt the lexical standard, but struggle to adopt the internal consistency standard which involves integrating and updating information into the situation model.

Chapters 4 – 6 report experiments that utilised self-paced reading time methodology. In comparison to off-line measures, self-paced reading time has the advantage of providing an insight into the comprehension monitoring processes that occur during reading and whether or not encountering erroneous information influences reading behaviour. In addition, to gain a greater insight into reading behaviour Chapter 3 reports an experiment that presented text non-cumulatively (i.e., text was presented one sentence at a time, each new sentence replaced the previous sentence), whereas Chapters 5 and 6 report experiments that presented text cumulatively (i.e., text was presented one sentence at a time, each sentence remained visible to the reader).

Chapter 4 explores how task instructions that do and do not alert children to the presence of errors in texts influence monitoring of general knowledge violations and internal inconsistencies. Note that nonwords were not included in this chapter or future chapters because performance was found to be near ceiling in Chapter 2. Children completed two conditions with different instructions. In the first, children were not explicitly alerted to the presence of errors in texts. In the second, children were alerted to the presence of errors in texts. Children's responses to the question 'did the story make sense?' and self-paced reading times were recorded. Findings were convergent with those found in Chapter 3: developmental differences and superior error detection for general knowledge violations over internal inconsistencies were found. Novel findings from the reading time measure reveal that task instructions that alert children to the likely presence of errors in texts change children's reading behaviour. Children spent more time reading texts when alerted to the presence of errors in texts. One likely reason for this is that children undertake a more purposeful and careful reading of the text. Children may slow down their processing to foster higher levels of evaluation and engage in critical monitoring strategies such as double-checking the sense of sentences. Because findings suggest that task instructions influenced children's monitoring behaviour, experiments included in further chapters in this thesis provided children with instructions that alerted them to the likely presence of errors in text and that their task was to evaluate the sense of information.

Chapters 5 and 6 explore general knowledge violations and internal inconsistencies, respectively. Within these chapters the presentation of text was also manipulated. Previous research has suggested that the visual presentation of text across the reading surface can influence a reader's ability to integrate information and establish text coherence (e.g., Just, Carpenter & Woolley, 1982). For this reason, in the first experiment included in each chapter each sentence of the text was presented on a separate line of the computer screen. Whereas, in the second experiment included in each chapter each sentence of the text was presented in connected text from left to right on each line of the computer screen.

Chapter 5 explores children's monitoring of general knowledge violations. The explicitness of the information forming the general knowledge violation was manipulated. Previous research has shown that children's monitoring of information can be influenced by the explicitness of the information that forms the error (e.g., Joseph, Liversedge, Blythe, White, Gathercole & Rayner, 2008; Markman, 1979). However, research is restricted to internal consistencies or sentences of text. In this work information was either related (e.g., using a knife rather than an axe to chop wood) or unrelated (e.g., using a fork rather than an axe to chop wood) to the standard state of affairs in terms of world knowledge. Children's responses to the question 'did the story make sense?' and self-paced reading times were recorded. The presentation of text did not influence children's performance on the sense question nor their real-time reading behaviour. Developmental difference in response to the sense question were found: older children made more correct sense judgements than younger children. Children made more correct sense judgements for unrelated than related general knowledge violations.

In contrast to the sense question findings, children's reading times for the target sentence were similar for related and unrelated general knowledge violations. Thus, the explicitness of error information influenced children's performance on the sense question, but not their real-time reading behaviour. This finding suggests that children may use the explicitness of information as a criterion for monitoring their comprehension, particularly when judging the sense of information.

Chapter 6 explores children's monitoring of internal inconsistencies. The distance between inconsistent information was manipulated. Previous research has shown that the distance between inconsistent information can influence children's error detection (e.g., Oakhill et al., 2005; van der Schoot, Reijntjes & van Lieshout, 2011). It has been suggested that such problems derive from the working memory demands involved in the integration of the two pieces of inconsistent information (e.g., Albrecht & O'Brien, 1993). In this work, inconsistent information was either adjacent (no intervening filler text), near (separated by one filler sentence) or far (separated by three filler sentences). Children's responses to the question 'did the story make sense?' and self-paced reading times were recorded. As found in the experiments included in Chapter 5, the presentation of text did not influence children's performance on the sense question nor their real-time reading behaviour. Developmental differences in response to the sense question found: older children made more correct sense judgements than younger children. Children made a similar number of correct sense judgements for consistent and adjacent internal inconsistencies, and more correct sense judgements for adjacent than far internal inconsistencies. In contrast to the sense question findings, children's reading times for the target sentence were similar for adjacent, near and far internal inconsistencies. Thus, the distance between inconsistent information influenced children's performance on the sense question, but not their real-time reading behaviour. This finding suggests that when inconsistent information is placed in adjacent sentences, with no intervening filler text, children find it easier to judge whether or not information makes sense than when information is separated by filler text. This may be because of the increased storage and processing demands associated with activating information from the evolving situation model and integrating information over several sentences.

Chapter 7 reports an experiment that utilised eye-tracking methodology, which offers advantages over both off-line and self-paced reading time measures. Critically eye-movements allow the distinction between initial reading and re-reading patterns. This means that it is possible to determine when errors inserted into a text first influence processing and therefore gain an insight into the time course of processing error information during comprehension monitoring. Eye-movements also provide an insight into whether or not children engage in monitoring strategies beyond reading the error sentence more slowly, such as looking back and re-reading text presented earlier in the passage.

Chapter 7 explores children's eye-movements when reading passages that contain internal inconsistencies. As in Chapter 7, the distance between inconsistent information was manipulated. Developmental differences were found: older children made more correct sense judgements than younger children. Children made more correct sense judgements for consistent than internal inconsistency passages, however the distance between inconsistent information did not influence performance. The eye-movement data revealed that younger children spent a similar amount of time reading the target region of the different passages types. This general increase in reading time for the target region may suggest that young children need to undertake a more purposeful and careful reading of texts that may contain internal inconsistencies in order to correctly identify whether or not they make sense. This finding may reflect young children's difficulty in monitoring their comprehension for internal consistency. In contrast, older children only spent more time reading the target region when the distance between inconsistent information was increased in the near and far manipulations. This finding suggests that older children adjust their reading behaviour depending on the distance between inconsistent information. This may be due to the increased storage and processing demands associated with integrating information across several sentences and evaluating the sense of this information. Further, surprisingly children were unlikely to go back and re-read the first piece of inconsistent information. It may be the case that children do not use, or are unaware of, this strategy of looking back to information presented earlier in a text. Alternatively, it may be the case that the type of contradictory information used to create internal inconsistency errors did not compel children to look back and reinspect information.

There were some general trends in findings across the experiments included in Chapters 3 – 7 which are discussed below. In addition to the primary aims stated above, the experiments included in these chapters investigated the relationship between comprehension monitoring and word reading, reading comprehension and working memory. Children completed measures of word reading (Test of Word Reading Efficiency, TOWRE; Torgesen, Wagner & Rashotte, 1999), reading comprehension (York Assessment of Reading Comprehension, YARC; Snowling et al., 2009) and working memory (sentence span and digit span tasks). In Chapter 3, only a sub-sample of children with the very best and the poorest comprehension monitoring completed the independent measures. Whereas, in Chapters 4 – 7 all of the children that participated in the experiments completed the independent measures.

Word reading was related to skill in monitoring general knowledge violations in younger children. It is likely that word reading accuracy is important for comprehension monitoring because reading fluency frees up processing resources that can be devoted to undertaking monitoring behaviours. In addition, reading comprehension was related to skill in monitoring internal inconsistencies. However, given the role that comprehension monitoring is thought to play in constructing a consistent and coherent representation of the text, this relationship was surprisingly weak. It is possible that this finding can be accounted for by limitations in statistical sensitivity that result from the relatively small sample sizes included in the independent measures analyses. Further, findings provided some evidence that both verbal and numerical working memory were related to monitoring of internal inconsistencies, particularly inconsistencies that were separated by intervening filler text. However, the strength of correlations were weak and no clear pattern of findings were found. Notwithstanding the limitations noted above, interestingly this finding questions the importance of working memory capacity as a source of monitoring difficulties. This suggests that comprehension monitoring difficulties may not result from failures in updating information into the situation model. Rather it seems that comprehension monitoring difficulties may arise from failures in constructing a richly elaborated situation model; children fail to include situation-relevant information that is required to interpret information included later in the text.

Chapter 8 provides the conclusions of the experimental work. First, a summary of the experimental findings related to developmental differences in comprehension monitoring are provided. Findings associated with task characteristics; error type and instruction, and reader characteristics; word reading, reading comprehension and working memory are discussed. Findings associated with the inclusion of off-line and real-time measures are also discussed. Second, limitations of the experimental work are discussed. Limitations focus on restrictions based on the sample, weakness of the error information, potential bias created by task instructions and suitability of the independent measures. Third, implications of the experimental work are discussed. Implications related to our understanding of comprehension monitoring research are discussed, including developmental differences in monitoring, the nature of the monitoring process, children's use of standards of evaluation and monitoring strategies, and sources of monitoring difficulties. Also, implications for education are discussed, of particular note is the importance of teachers including instructional support that reminds children to construct a richly elaborated situation model and check their understanding during reading. Finally, implications for future experimental work are discussed, for example exploring the complexity of text content as a source of monitoring difficulty and exploring the relationship between monitoring and other skills such as attentional control utilising larger sample sizes.

In conclusion, the data demonstrate developmental differences in comprehension monitoring. Older children were shown to be better at correctly judging the sense of information and more likely to adjust their reading behaviour in relation to error information. It seems that both age groups undertake similar monitoring behaviours, albeit with different levels of success. A range of factors were shown to influence monitoring skill, including the type of error inserted into texts. Children were most able to adopt the lexical standard and least able to adopt the internal consistency standard. There was some evidence that within error manipulations, such as the explicitness of error information and the distance between inconsistent information, influenced children's ability to accurately judge whether or not information makes sense. Also, the extent to which task instructions alerted participants that texts contained errors influenced reading behaviour. Children undertook a more purposeful and careful reading of the text when alerted that texts may contain errors. Furthermore, the data reveals that the relationship between comprehension monitoring and working memory capacity is relatively weak. Thus, suggesting that monitoring difficulties may arise from failures in constructing a richly elaborated situation model, rather than failures in updating the situation model. In the future, longitudinal work that includes a broader age range and larger sample sizes is needed to extend findings and provide more precise information on the developmental trajectory of comprehension monitoring, and the nature of the relationship between monitoring ability and critical component skills.

CHAPTER TWO

INTRODUCTION

Comprehension monitoring, also referred to as self-regulated reading (e.g., Hacker, 1998) and validation (e.g., Singer, 2013), is a metacognitive skill involving the operation of two component processes: evaluation and regulation (e.g., Baker, 1985). Evaluation refers to a reader's assessment of their current state of comprehension while reading. Regulation occurs when a reader has evaluated their understanding, found it inadequate, and consequently selected and implemented a strategy, such as re-reading the text, to restore comprehension. These processes can be viewed as executive functions, directing a reader's cognitive processes as they strive to make sense of incoming textual information (e.g., Wagoner, 1983).

Comprehension monitoring has been the focus of cognitive, social and educational psychology. The ability to detect inconsistencies in texts, to understand that inconsistencies impair comprehension, and to take remedial action to restore comprehension are clearly important skills for a reader. This thesis explores the development of comprehension monitoring (both evaluation and regulation components) in children aged 7 to 10 years, and the critical task and reader characteristics that influence this skill.

2.1. Comprehension monitoring and reading comprehension

Comprehension is often defined as the construction of a mental representation (i.e., a situation model) of the situation described in a text, including both textual information and general world knowledge (e.g., Kintsch, 1998). Whereas, comprehension monitoring is often defined as the monitoring of information consistency during comprehension (e.g., Singer, 2013). Comprehension monitoring is an important component skill of comprehension (e.g., Cain, Oakhill & Bryant, 2004; Rubman & Waters, 2000). This is because successful comprehenders build a situation model by monitoring a number of key situational dimensions, for example:

time, space, causation, and the protagonists' characteristics, goals and emotions (e.g., Zwaan, Langston & Graesser, 1995). By integrating information from these dimensions, and incorporating new sentences or clauses into the evolving model, readers gradually update their representation and build a coherent and richly connected situation model (e.g., van der Schoot, Reijntes & van Lieshout, 2011). Working memory plays an important role in these integration and updating processes (e.g., Radvansky & Copeland, 2001). Working memory is a limited capacity system for the simultaneous storage and processing of information (Baddeley & Hitch, 1974). When integrating incoming information into the evolving situation model readers are required to activate prior information from their situation model, which is assumed to be stored in long-term working memory. When the situation model must be updated, readers are required to keep newly encoded information active in working memory while it adds to or replaces old information. It is suggested that the better these processes, the more consistent and coherent the construction of the situation model (e.g., Radvansky & Copeland, 2001).

Despite the importance of comprehension monitoring for comprehension, comprehension monitoring is not often explicitly discussed in current models and theories of reading comprehension (e.g., Construction-Integration Model (Kintsch & van Dijk, 1978); Landscape Model (Tzeng, van den Broek, Kendeou & Lee, 2005); Structure Building Model (Gernsbacher, 1990); Resonance Model (Albrecht & O'Brien, 1993); Event-Indexing Model (Zwaan, Langston & Graesser, 1995); Casual Network Model (Trabasso, van den Broek & Suh, 1989); Constructionist Model (Graesser, Singer & Trabasso, 1994)). One model that perhaps has the strongest assumptions regarding comprehension monitoring processes during comprehension is the Construction-Integration Model (Kintsch & van Dijk, 1978). According to the model, a construction phase involves the activation of a network of interconnected ideas. This is followed by an integration phase in which activation stabilises and the most active ideas are maintained in the representation, and the weaker ideas are removed from the representation.

Because general world knowledge plays an important role in stabilising information, the process is, by definition, influenced by comprehension monitoring.

Further, the nature of comprehension monitoring processes remain largely unaddressed (e.g., Kendeou, 2014). Two opposing views dominate the literature. One view is that comprehension monitoring is perceived as strategic and optional process that requires an ‘evaluative mindset’ (e.g., Gilbert, 1991; Herbert & Kubler, 2011; Wiswede, Koranyi, Muller, Langner & Rothermund, 2012). An alternative view is that comprehension monitoring is a passive and non-optional processes that does not require an ‘evaluative mindset’ (e.g., Isberner & Richter, 2013; Singer, 2006). Based on this view, information during comprehension is routinely evaluated, and monitoring skill develops with age and experience (e.g., Singer, 2013).

In recent years, an attempt to explicitly situate comprehension monitoring in the context of current reading comprehension models has been presented by Cook and O’Brien (2014). Their RI-Val view of comprehension posits that comprehension monitoring is a process that operates in parallel with the activation and integration stages of processing. That is, when information is activated, it is integrated with the contents of working memory and validated against general world knowledge and earlier portions of the text representation.

In summary, comprehension monitoring is a higher level language skill that plays an important role in reading comprehension. It is noteworthy that comprehension monitoring is not often explicitly discussed in current models and theories of reading comprehension, even though most models implicitly assume its presence and importance. Further research in the field is warranted to comprehensively establish the nature of the comprehension monitoring process and its role in reading comprehension.

2.2. The design and aims of the research

Many investigations into comprehension monitoring have focused their attention on the effects of reading and comprehension ability on children's evaluation of comprehension. In addition, the findings from studies that have explored comprehension monitoring developmentally are inconsistent. Thus, we have limited information regarding the developmental changes in children's comprehension monitoring. The population of particular interest to this thesis were 7- to 8-year-old and 9- to 10-year-old children. These age groups were chosen as representative of the period of transition between beginning and fluent reading in children in British schools (e.g., Nation & Snowling, 1998). Thus, the experiments in this thesis were designed to build on the comprehension monitoring literature, in order to further investigate the skills and processes that are important in the development of comprehension monitoring. This was achieved by comparing the performance of younger children to older children, which provided an insight into differences in monitoring skills between the two age groups.

The experiments included in this thesis focus on comprehension monitoring of text. This is because in contrast to listening comprehension, reading comprehension provides the reader with greater control over the rate of input. That is, readers have the opportunity to omit portions of text, re-read sections and pause on particular words, phrases or sentences (e.g., Just & Carpenter, 1980). Thus, a reader can take in information at a pace that matches their internal comprehension processes.

Many investigations into comprehension monitoring focus on children's detection of internal inconsistency errors. However, few investigators have treated comprehension monitoring as a multidimensional process by examining children's ability to evaluate their understanding at different levels, and consequently use various standards of evaluation. Standards of evaluation refer to a set of implicit or explicit criteria for comprehension monitoring (e.g., Baker, 1984a), for example, lexical, external consistency and internal consistency. To address this, the work reported in this thesis included tasks that explored monitoring of nonwords, general knowledge violations and internal inconsistencies. Furthermore, research has shown that children experience difficulty in a range of comprehension monitoring situations. The work reported in this thesis involved further examination of some of these key areas, namely task instructions and readers' working memory capacity.

2.3. Literature review

Several different factors related to task and reader characteristics have been hypothesised to influence children's comprehension monitoring. The aim of this chapter is to provide an overview of the theory and research in this area as background to the experimental work conducted in this thesis. First, task characteristics, specifically the type of error inserted into texts and the nature of the instructions given to participants are considered. Second, reader characteristics, specifically knowledge and use of reading goals, standards of evaluation, and working memory capacity are considered. Finally, ways to assess comprehension monitoring, both off-line and in real-time, are considered.

2.3.1. Task characteristics

In this section characteristics of the task that may influence comprehension monitoring are examined. The characteristics reviewed here are the type of error a participant must detect and the nature of the instructions given to participants.

Error type

In practice, comprehension monitoring often comes down to detecting and, if possible resolving inconsistencies such as information conflicting world knowledge and contradictory sentences. Therefore, research investigating comprehension monitoring has typically used the error detection paradigm. This paradigm involves introducing the participant to material that contains experimenter-inserted errors. Participants' ability to monitor their comprehension is measured by asking them whether or not the material made sense, and/or to identify any 'bits that do not make sense.' The assumption underlying the error detection paradigm is that the inserted errors disrupt comprehension, so readers who are effectively monitoring their ongoing comprehension should be able to detect and identify these errors. Although children may not typically be required to detect experimenter inserted errors in their evaluation of text during day-to-day reading, the error detection paradigm has been shown to clearly tap comprehension monitoring skills (e.g., Ruffman, 1996).

Readers' response to several different error types has been studied. This thesis will focus on the three most prominent error types within the comprehension monitoring literature: nonwords, general knowledge violations and internal inconsistencies. Successful monitoring of these error types requires the reader to adopt different standards of evaluation. This thesis focuses on the lexical, external consistency and internal consistency standards, which correspond to the monitoring of nonword, general knowledge violation and internal inconsistency errors, respectively (e.g., Baker, 1985). The lexical standard of evaluation requires readers to check that a word is present in their mental lexicon. The external consistency standard requires readers to compare the information presented with their prior knowledge. The internal consistency standard requires readers to maintain or access a memory representation of previously presented text, and evaluate its consistency with just-read information. The application of these standards imposes different information processing demands on the reader, which are discussed for each error type below. In addition, nonword, general knowledge violation and internal inconsistency errors relate differently to the construction of the situation model. The way that different error types relate to the construction of the situation model is discussed for each error type below.

Nonword

In research that involves nonword errors, experimenters replace a single word in a passage with a string of letters that is not a real word. An example is (nonword shown in italics): 'It was so hot that most *brugens* would melt there' (Baker, 1984a, p.293). Although nonword errors typically look (or sound) like a word, they are different from new words that young children encounter when reading. This is because nonword errors violate the lexical principles accepted by a native speaker of a language. Thus, nonword errors provide researchers with information about whether children use lexicality as a criterion for monitoring their comprehension.

Research has demonstrated that in comparison to general knowledge violation and internal inconsistency errors, children are better at monitoring their comprehension for nonword errors (e.g., Baker, 1984a; Garner, 1981; Markman & Gorin, 1981). This may be because children's spontaneous reading goals focus on decoding, so word level errors are more easily detected (e.g., Myers & Paris, 1978). In addition, depending on the kind and amount of processing needed to uncover the error, children's sensitivity to the error will vary (e.g., Markman & Gorin, 1981). In comparison to general knowledge violation and internal inconsistency errors, the working memory demands associated with monitoring nonword errors are relatively low; readers are only required to check whether or not a word is in their mental lexicon (e.g., Baker, 1984a; Oakhill, Hartt & Samols, 2005). Also, this monitoring process can be considered to be relatively independent from the construction of the situation model. This is because the process of checking one's mental lexicon requires limited interaction with the situation model in terms of activating, integrating or updating information from the model.

Despite the comparative ease with which 9- to 11-year-old children identify nonwords, research has shown that monitoring of this error type is not perfect in children of a similar age. For example, Paris and Myers (1981) investigated 9- and 10-year-old good and poor readers monitoring of nonwords and nonsense phrases. Poor readers were less likely to detect nonword errors, and even 10-year-old good readers identified fewer than half of the inserted nonwords. Thus, good and poor readers, as well as older children, experience problems in monitoring their comprehension when assessed using nonwords. However, a major limitation of this study is that the materials were more difficult for the poor readers than the good readers. Consequently, the poor readers' ability to monitor their comprehension for nonword errors was confounded with their ability to read the stories.

This limitation has been overcome by Oakhill et al. (2005) who used the procedure of Paris and Myers (1981) to investigate the differences in comprehension monitoring skills of 9- to 10-year-old good and poor comprehenders, matched for reading vocabulary and word recognition skills. They found that there was no significant difference between good and poor comprehenders in the number of nonwords correctly identified. Thus, when controlling for vocabulary and word reading skills it seems that good and poor comprehenders are equally able to monitor their comprehension at the word-level, and subsequently detect and identify a reasonable number of nonword errors.

General knowledge violation

General knowledge violation errors are created by inserting information that violates general knowledge into a passage. An example is (general knowledge violation shown in italics): 'They used *sand from the trees* to make many things' (Baker, 1984a, p.293). To detect a general knowledge violation, a proposition from the text must be held in working memory and compared with stable information from the readers' knowledge base (long-term memory). Similarly to nonword errors, this monitoring process involves comparing information from the text with information held in long-term memory, and it can be considered to be relatively independent from the integration and updating processes involved in the construction of the situation model.

Research has demonstrated that older children are better than younger children at monitoring their comprehension for external consistency. For example, Baker (1984a) investigated 5- to 11-year-old children's monitoring of a number of errors including general knowledge violations. Children were asked to listen to short stories and tell the experimenter about any mistakes that they detected. Baker (1984a) found that older children reported more general knowledge violations to the experimenter than younger children. Also, older children were more likely to report general knowledge violations without prompting from the experimenter. However, these results must be interpreted cautiously because more than one error and different error types were inserted into each passage. It may be that once a child had detected one error within a passage they considered the task complete and did not continue to monitor the remainder of the text for errors. Also, once a child had detected an error using one standard of evaluation they may not have monitored the remainder of the text using different standards.

Research has also demonstrated differences in the comprehension monitoring abilities of high performing students compared to low performing students in response to general knowledge violation errors. For example, Owings, Petersen, Bransford, Morris and Stein (1980) asked the least and most successful 9- to 11-year-old students within a class (based on the judgements of two class teachers that were corroborated by a State test of basic skills) to read and study stories that varied in the degree to which they made sense relative to participants' prior general knowledge. The more successful students appeared to spontaneously monitor for external consistency as they read and studied the stories, because they were better able to evaluate the extent to which new information made sense relative to their prior knowledge. They were also able to accurately answer memory questions about stories containing a greater degree of conflict with their prior knowledge. However, less successful students only showed awareness of comprehension difficulty when prompted to evaluate their comprehension and identify which passages made sense. Thus, research suggests that in comparison to low performing students, high performing students are better able to monitor their comprehension in terms of consistency with prior general knowledge.

Internal inconsistency

Experimenters create internal inconsistency errors by making information in the passage conflict with previously presented information, for example (inconsistent information is shown in italics): 'The *temperature on Venus is much higher than boiling water*…. But *it is much too cold for us to live there*' (Baker, 1984a, p.293-294). Much of the research on comprehension monitoring focuses on readers' detection of internal inconsistency errors. It is likely that this is because internal inconsistency errors provide a greater insight into the construction of the situation model and broader comprehension processes. According to the situation model theory (Zwaan & Radvansky, 1998), inconsistencies specifically impede the updating process, during which a new sentence or clause is incorporated into the evolving situation model. Unsurprisingly, new information is more difficult to integrate into the evolving model if it is inconsistent with the information in the current state of the model (e.g., van der Schoot et al., 2011).

Research has shown that young children and poor comprehenders are particularly poor at monitoring their comprehension for internal inconsistency errors (e.g., Baker, 1984a; Oakhill et al., 2005; van der Schoot et al., 2011). An early example comes from Markman (1979). She investigated the limits of 8- to 9-year-old and 11- to 12-year-old children's spontaneous comprehension monitoring ability. Children listened to essays containing explicit or implicit internal inconsistencies and responded to a series of probe questions. Findings reveal that children identified internal inconsistencies on fewer than half of all occasions. Also, children in both age groups failed to ask questions that challenged the comprehensibility of essays. Thus, there were no differences in children's monitoring of internal consistency between the age groups studied. Markman (1979) concluded that children in all age groups seemed genuinely unaware of the presence of internal inconsistency errors. Thereby suggesting that children have problems monitoring their comprehension for internal consistency. However, Pratt and Nesdale (1980) have suggested that the findings of Markman's (1979) work may be limited by the content of the essays in which internal inconsistencies were inserted. The majority of errors concerned animal behaviour, and therefore information with which the children may not have been familiar. Consequently, children may have overlooked inconsistencies or delayed judgement of this information until subsequent information clarified the confusion.

In contrast to Markman's (1979) work, research has shown that children of a similar age can successfully monitor their comprehension for internal consistency. For example, van der Schoot et al. (2011) used a reading time paradigm to investigate 10- to 12-year-old good and poor comprehenders' monitoring of internal inconsistencies. Children read stories in which the action of the protagonist was either consistent or inconsistent with a description of the protagonist's character (e.g., a fast-food addict eating a hamburger vs. a vegetarian eating a hamburger). The distance between inconsistent information was manipulated (adjacent vs. separated by a long filler paragraph). The purpose of the filler paragraph was to ensure that the contradictory information was not concurrently active in working memory. To detect inconsistencies in the distant condition, readers need to activate prior information from their situation model, which is assumed to be stored in long-term working memory, before they can check its consistency with new incoming information (e.g., Ericsson & Kintsch, 1995; Was & Woltz, 2007). The greater the distance between the inconsistent information, the greater the storage and processing demands involved in the co-activation of information.

Good comprehenders had longer reading times for the inconsistent than consistent target sentences in both the adjacent and separated conditions. However, poor comprehenders only showed this pattern in the adjacent condition. This suggests that children were able to successfully monitor their comprehension for internal consistency when contradictory information was adjacent. However, when contradictory information was separated by filler text, poor comprehenders struggled to monitor their comprehension for internal consistency. Thus, suggesting that the distance between contradictory information influences children's, particularly poor comprehenders', detection of internal inconsistency errors.

In summary, these studies demonstrate that developmental differences exist in children's comprehension monitoring: Older and better readers are more able to monitor their comprehension than younger and poorer readers. In general, children are better at monitoring their comprehension for nonword errors than general knowledge violation and internal inconsistency errors. However, findings across studies are inconsistent, and there is a lack of robust research using a range of error types that explores comprehension monitoring developmentally. Also, there are a number of methodological weaknesses in the existing research (e.g., including multiple error types in one passage). To address these issues and extend our knowledge of the development of comprehension monitoring skills, the main body of work included in this thesis investigates 7 to 10-year-old children's monitoring of nonword, general knowledge violation and internal inconsistency errors.

Instructions

The instructions an experimenter provides a participant with influences the manner in which materials are processed, and consequently comprehension monitoring behaviours (e.g., Baker, 1984a; Markman & Gorin, 1981; Zabrucky & Moore, 1989; Winograd & Johnston, 1982). Grabe, Antes, Thorson and Kahn (1987) suggest that comprehension monitoring research can be framed within a continuum that describes the extent to which participants are aware of their involvement in a study assessing their ability to detect and identify errors. This thesis focuses on studies that compare the performance of participants that are alerted that materials may contain errors and their task will involve the detection of these errors, with participants that are not alerted that materials may contains errors. Such a comparison provides an insight into children's spontaneous comprehension monitoring, and the potential benefits in monitoring that alerting participants to the presence of errors in materials may provide.

It has been suggested that it is important to provide participants with instructions that alert them to the presence of errors in materials because in the absence of this information readers have been shown to undertake effortful resolution processes (e.g., Baker, 1979), for example, drawing upon general knowledge to render the text more comprehensible. Such processes may impact upon error identification because the error may become resolved in the reader's mind, so it is no longer perceived as problematic (e.g., Garner, 1981). Also, the cognitive effort involved in undertaking resolution processes may compromise the reader's memory representation of the text (e.g., Ackerman, 1984a). Thus, it may be argued that alerting participants to the presence of errors in materials is essential if a true reflection of their ability to monitor their comprehension is to be gained (e.g., Baker, 1985; Zabrucky & Moore, 1989). However, it should be noted that the findings of studies that alert participants to the presence of errors in texts should take into consideration the possibility that such instructions might lead to changes in reading behaviour (e.g., Grabe et al., 1987). For example, participants may undertake tasks with greater care and attention, which could result in more monitoring behaviours and consequently increased error detection.

Markman's (1979) work investigating 8- to 11-year-old children's monitoring of explicit and implicit internal inconsistencies is discussed above in the internal inconsistency section. In an extension of this work, Markman (1979) manipulated the instructions children received. Children were either asked to determine the comprehensibility of essays or informed that essays contained a problem that they should try to detect. A significant effect of instruction was found both within age groups and developmentally. Children receiving instructions that alerted them to the presence of errors performed significantly better than children instructed to determine the comprehensibility of essays. Also, overall, implicit inconsistencies were more difficult than explicit inconsistencies, however, for 11- to 12-year-olds this difference disappeared when children received instructions that the essays contained a problem. Thus, it seems that children are better able to detect and identify errors in materials when they are forewarned that problems are present. Additionally, if older children are provided with instructions that alert them to the presence of inserted errors within materials, they are capable of carrying out the integrative processes necessary to detect and identify even complex error types, such as implicit internal inconsistencies.

Work investigating the influence of alerted versus non-alerted task instructions on children's comprehension monitoring is presented in Chapter 3 of this thesis. Other work in this thesis provides children with instructions that alert them to the presence of errors in text, so that children's monitoring behaviours when they are aware of the task to detect errors can be explored further.

2.3.2. Reader characteristics

In this section characteristics of the reader that may influence comprehension monitoring are examined. These characteristics include children's knowledge and use of different reading goals, standards of evaluation and working memory capacity.

Reading goals

One reason suggested for developmental differences in comprehension monitoring is that younger children have less knowledge of their own cognition and their use of cognitive resources than older children (i.e., metacognition) (e.g., Ehrlich, 1996). A reader's knowledge about comprehension processes and their use of reading goals and strategies to support comprehension will influence their approach to different reading situations (e.g., Ruffman, 1996), and consequently their ability to monitor their comprehension. Such metacognitive knowledge and skills develop during childhood (e.g., Baker & Brown, 1984).

Research investigating children's understanding of text suggests that younger readers set different reading goals to older children (e.g., Cross & Paris, 1988, Myers & Paris, 1978). Reading goals can either be spontaneously set by the reader or externally set through instructions provided by the experimenter. In this section, the focus is on children's spontaneous reading goals. Younger readers appear to have limited knowledge about reading, and they tend to perceive reading primarily as a decoding task. In comparison, older readers approach reading with a different criteria, focusing on extracting meaning from text. Myers and Paris (1978) found evidence of this when assessing 8- and 12-year-olds' reading knowledge. When asked how they would retell a story to a friend, younger children favoured using the exact wording of a story, whereas older children favoured using the gist. Further, when asked how they would deal with unknown words younger children favoured 'sounding-out' strategies, rather than 'meaning-related' strategies.

Research has also shown that young readers do not actively include constructive processing in their reading goals (e.g., Markman, 1977, 1979). Constructive processing reflects readers' efforts to build a situation model of the text by connecting individual propositions together during reading (e.g., Schmidt, Schmidt & Tomalis, 1984). Rather than integrating and comparing the comprehensibility of propositional information at the text-level, young readers process information at the word- or sentence-level (i.e., word-by-word or sentence-by-sentence) (e.g., Canney & Winograd, 1979; Garner, 1981; Markman, 1977, 1979). Thus, young readers fail to construct a coherent and richly connected representation of the text, which can be evaluated for consistency. Markman (1977) found evidence of this when exploring 6- to 8-year-old children's ability to determine the adequacy of instructions for playing a game or carrying out a magic trick. Despite instructions missing key information, younger children failed to question their adequacy. However, when experimenter led demonstrations accompanied the instructions younger children were more likely to notice the inadequacy of the instructions. Because demonstrations reduced processing demands, it was suggested that children's initial insensitivity to the instructions inadequacy was due to a lack of constructive processing. Note that the reasons why readers may fail to engage in constructive processing are not limited to reading goals. Additional reasons for failure to engage in constructive processing include limitations in reading fluency, general comprehension ability and working memory capacity (e.g., Oakhill, 1982).

If young readers have less knowledge about different ways of reading for different goals and fewer strategies for repairing comprehension failure, then their ability to monitor their comprehension may be limited and may only occur in certain situations (e.g., when task instructions provide participants with a reading goal that alerts to the presence of errors). In addition, if young children's reading goals focus on checking that words make sense and processing information at the word- and sentence-level, it is unsurprising that they fail to monitor their comprehension for more global, text-level errors. In the work included in this thesis, findings are related to reading goals.

Standards of evaluation

A further reason posited for developmental differences in comprehension monitoring is that younger children are less able to adopt different standards of evaluation (e.g., Baker, 1984a, 1984b). These standards influence readers' evaluation of text, and so impact upon their ability to monitor their comprehension (e.g., Baker, 1984a, 1984b). Research suggests that some standards are more likely to be adopted by young children than others. Baker (1984a) found that 5- to 11-year-old children were more likely to spontaneously adopt a lexical standard, than an external or internal consistency standard. In addition, Markman and colleagues (e.g., Markman, 1976, 1978, 1979; Markman & Gorin, 1981; Osherson & Markman, 1975) found that, in general, young readers judge their comprehension of a text by evaluating the truth of individual statements, rather than relating propositions within the text to check their internal consistency (i.e., applying an external consistency, rather than an internal consistency standard). Such differences in the application of standards may be due to children's developing knowledge of different standards, and the ease with which different standards can be applied (e.g., Baker, 1984a, 1984b). The application of lexical, external consistency and internal consistency standards impose different information processing demands, and consequently children's ability to adopt these standards may depend on the kind and amount of processing required (e.g., Baker, 1985). As discussed above in the Error Type Section (2.3.1), the working memory demands associated with adopting the lexical and external consistency standards are not as demanding as those associated with the internal consistency standard. This is because the lexical and external consistency standards require readers to evaluate text at the word- and sentence-level, whereas the internal consistency standard requires readers to integrate and compare information at the text-level.

Readers' standards of evaluation may also include consideration of the strength of coherence that they aim to maintain for the text. This is reflected in van den Broek's standards of coherence (van den Broek, Risden & Husebye-Hartmann, 1995). It is possible that young children adopt standards that include low coherence (e.g., Glenberg, Wilkinson & Epstein, 1982). Thus, children may fail to monitor their comprehension because their criteria for what constitutes erroneous information is low, and they do not consider the types of errors included in monitoring tasks as problematic. Also, children who adopt standards that include low levels of coherence may detect errors, but assume that subsequent information will resolve the problem (e.g., Baker, 1979). Thus, children may fail to give appropriate weight to erroneous information when evaluating the sense of information and constructing their situation model.

Conversely, children may adopt standards that include adequate coherence and result in error detection, but they may use 'fix-up' strategies, such as inferences or ignoring erroneous information, to try to restore comprehension (e.g., August, Flavell & Clift, 1984). Thus, even if the reader does not get the author's intent meaning, they feel that they have an adequate understanding of the text. Garner (1981) found evidence of this when exploring 10- to 12-year-old good and poor comprehenders ratings of passage comprehensibility. Children were asked to read, rate, and indicate why passages rated as incomprehensible were difficult. In general, children rated erroneous messages as 'okay', rather than 'difficult to understand'. However, findings should be interpreted with caution because children only rated the comprehensibility of three passages (one per experimental condition). Work included in this thesis relates findings to standards of evaluation.

Working memory

A different reason suggested for developmental differences in comprehension monitoring ability is that younger children are less efficient at simultaneously storing and processing information in working memory (e.g., Seigneuric, Ehrlich, Oakhill & Yuill, 2000). Throughout childhood children's working memory capacity increases (e.g., Gathercole, 1998), and children become more efficient at undertaking basic reading tasks, such as decoding (e.g., Vosniadou, Pearson & Rogers, 1988). This means that older children have access to greater processing resources, which can be devoted to comprehension monitoring (e.g., Perfetti & McCutchn, 1987).

Whilst it is well established that there is a strong relation between working memory and children's reading comprehension (e.g., Cain et al., 2004; Oakhill et al., 2003; Seigneuric et al., 2000; Yuill et al., 1989), the relationship between working memory and comprehension monitoring is less clear. Only a handful of studies investigating children's comprehension monitoring have included an independent assessment of working memory (e.g., Oakhill et al., 2005; van der Schoot et al., 2011), and in the main these studies have included one age group and focused on good and poor comprehenders. Despite these limitations, it has been suggested children's working memory is important for comprehension monitoring, particularly the monitoring of internal consistency (e.g., Oakhill et al., 2005; Perfetti et al., 1996; van der Schoot et al., 2011). This is because successful comprehension monitoring requires both the processing and storage of information. Readers must hold previously read information in memory, while new incoming information is encoded and integrated into their situation model (e.g., Yuill & Oakhill, 1991). For this reason, children who have poor working memory may be poor at monitoring their comprehension. It may be that these children can monitor their comprehension when the working memory demands of the task are low, for example detecting word- and sentence-level errors. But, comprehension monitoring breaks down when the working memory demands are high, for example detecting text-level errors that require the integration of information over several sentences (e.g., Oakhill, 1996).

The relation between comprehension monitoring and working memory has been investigated using working memory tasks that include sentences and numbers (e.g., Oakhill et al., 2005; van der Schoot et al., 2011). However, it is not yet clear which type of task is the most appropriate. Research exploring children's reading comprehension demonstrates that children's verbal working memory is more highly correlated with reading comprehension than numerical working memory (e.g., Cain et al., 2004; Oakhill, Yuill & Garnham, 2011; Seigneuric et al., 2000). This is unsurprising given that verbal working memory tasks are typically word- or sentence-based, so are demanding of linguistic skills. Whereas, numerical working memory tasks are typically number-based, so do not require the processing of words and sentences.

A crucial question that remains is whether individual differences in working memory capacity underlie individual differences in comprehension skills that require the integration of information, such as comprehension monitoring, or whether deficiencies in these skills exist in the presence of adequate working memory (and lexical processing) skills. To address this question and further explore the relationship between comprehension monitoring and working memory, independent measures of working memory were included in the experiments in this thesis. Both verbal and numerical measures were included to explore whether any relation is mediated by linguistic ability. In addition, experiments presented in Chapters 6 and 7 manipulated the working memory demands associated with detecting internal inconsistency errors.

2.3.3. Ways to assess comprehension monitoring

Comprehension monitoring is a multi-dimensional process that has been investigated using a variety of measures. In the sections that follow, off-line measures will be considered first, followed by real-time measures. Note that off-line and real-time measures are also referred to as product and process variables, respectively.

Off-line measures

Off-line measures are collected after the participant has finished reading. Examples include underlining of text (e.g., Oakhill et al., 2005), verbal reports of passage comprehensibility and responses to probe questions (e.g., 'did the story make sense?') (e.g., Baker, 1984a, 1984b).

There has been some debate about the phrasing of probe questions. Ruffman (1996) argues that asking children whether the story makes sense taps comprehension monitoring skills more clearly, than asking children whether there is anything wrong with the text. This is because 'making sense' means 'understanding' or 'comprehending'. Thus, children are not only required to detect errors in text, but also understand that the erroneous information impairs their comprehension.

The isolated use of off-line measures features prominently in the early comprehension monitoring literature, however such measures have some notable limitations. Because off-line measures are collected after reading, they only provide information about the end product of comprehension monitoring processes (e.g., Zabrucky & Ratner, 1986). That is, the readers understanding of the full text, which includes any regulation strategies undertaken to repair miscomprehension. Thus, off-line measures do not provide any information about ongoing comprehension monitoring processes that occur during reading. Also, off-line measures assume that errors were not detected if participants do not acknowledge the presence of these errors in responding to post-reading activities (e.g., Grabe et al., 1987). Consequently, off-line measures do not account for instances when readers' have detected errors during reading, but failed to report this monitoring when responding to post-reading activities.

In addition, off-line measures require participants to remember and verbalise their comprehension processes. The validity of verbal reports of cognitive processes (e.g., Nisbett & Wilson, 1977) and more specifically comprehension monitoring (e.g., Zabrucky & Ratner, 1986) has been questioned. It has been suggested that data from verbal report measures may be the result of uninteresting artefacts of experimental designs (e.g., an inability to recall erroneous information or a hesitancy to criticise the experimenter), rather than an accurate reflection of children's monitoring behaviours. Moreover, research suggests that verbal reports may reflect the monitoring behaviours of some age and reading ability groups less accurately than others (e.g., Zabrucky, Moore & Ratner, 1985).

Real-time measures

Real-time measures are collected during the reading process, so measure aspects of behaviour that are concurrent with ongoing comprehension monitoring processes. Examples include vocalisations of words (e.g., hesitations, repetitions and self-corrections) (e.g., Carpenter & Daneman, 1981), facial expressions (e.g., puzzlement and surprise in young infants) (e.g., Patterson, Cosgrove & O'Brien, 1980), reading times (e.g., Baker & Anderson, 1982), eye-movements (e.g., Daneman & Carpenter, 1983) and event-related potentials (ERPs; e.g., N400 and P600 effects) (e.g., Ferretti, Singer & Patterson, 2008). Although the vocalisations of words, facial expressions and ERPs have been shown to provide valuable insights into comprehension monitoring, these measures are beyond the scope of this thesis.

Real-time measures have several advantages. First, real-time measures provide an insight into comprehension monitoring processes as they unfold during reading. This means that researchers can undertake a more systematic exploration of the specifics of children's monitoring behaviours, for example, using reading time and eye-movement measures to explore the time spent reading (and thus comprehending) erroneous versus consistent information. Therefore, an insight into the processes underlying the development of children's comprehension monitoring skills may be obtained (e.g., Zabrucky & Ratner, 1986). Experimenters can also gain an insight into processes beyond comprehension monitoring, for example a reader's engagement in constructive processing.

Second, in comparison to off-line measures, real-time measures may be a more sensitive indicator of emerging comprehension monitoring, particularly in investigations of error detection (e.g., Brown, Bransford, Ferrara & Campione, 1983). For example, real-time measures sometimes suggest that children have detected an error, although they may not explicitly identify the error through off-line verbal report measures (e.g., Harris, Kruithof, Terwogt & Visser, 1981). Researchers have suggested a number of reasons why children may exhibit real-time evidence of comprehension monitoring in the absence of off-line evidence. One view is that children may have processed the text constructively and detected an error whilst reading, but did not or could not 'act on' or remember the internal detection signal generated during reading (e.g., Harris et al., 1981). It has also been suggested children are more likely to make a real-time response because such responses may be easier to produce because they do not require a verbal or physical response (e.g., Patterson et al., 1980).

Third, real-time measures can be used without requiring readers to purposefully search for errors, thus they are less likely to bias readers' behaviour (e.g., Grabe et al., 1991). Finally, in comparison to other measures available for investigating language processes that occur during reading, it can be argued that eye-movements provide a natural means of assessment and offer an insight into the moment-by-moment comprehension monitoring processes (e.g., Rayner, 1998). However, to-date there are a limited number of studies that have explored eye-movements during comprehension monitoring in children.

In summary, it is clear that both off-line and real-time measures demonstrate sensitivity to comprehension monitoring behaviours. However, these measures provide an insight into different aspects of comprehension monitoring processes: those that occur during and after the reading process. Thus, these measures may tap different aspects of comprehension monitoring, and demand different skills from participants. Consequently, one cannot assume that a particular pattern of behaviour while reading will predict whether or not the reader will state that an error has been identified. To account for the multi-dimensional nature of comprehension monitoring, the experimental work included in this thesis utilises both off-line and real-time measures.

2.4. Conclusions

This chapter has touched on a variety of task and reader characteristics, which may affect comprehension monitoring ability. The following is a summary of the main issues raised in this review and their relevance to the work conducted in this thesis.

The review of the literature highlighted developmental differences in children's monitoring of different error types. Older readers are generally better at detecting and identifying all error types, and this is particularly true for more complex errors such as internal inconsistencies. The detection of different error types requires the adoption of different standards of evaluation, which impose different processing demands. Difficulties may be caused by a lack of knowledge and experience utilising different standards or the inability to adopt different standards due to the associated processing demands. The main body of work included in this thesis focuses on investigating children's monitoring of nonword, general knowledge violation and internal inconsistency errors.

The instructions readers receive alerting them to the presence of errors in text and their task of monitoring their comprehension have been shown to influence monitoring behaviour. Studies have provided children with instructions that do and do not alert them to the presence of errors in texts and the task of detecting these errors. Findings suggest that children are able to demonstrate monitoring behaviours in response to these different instructions. However, it may be the case that when readers are not provided with instructions that alert them to the presence of errors findings do not accurately reflect monitoring skills. This is because readers tend to undertake 'fix-up' strategies to resolve problematic information when they are unaware that the information may contain errors. Thus, it is possible that readers may have detected errors, but resolved them. Also, engaging in such processing may limit the cognitive resources available for monitoring. Consequently, it can be argued that to gain a true insight into comprehension monitoring skills participants must be provided with instructions that alert them to the presence of errors in texts and their task of detecting errors. Only when participants are provided with such instructions can the experimenter be sure that the participants' initial reading goals are in line with their expectations. Work included in this thesis further explored the influence of task instructions on children's comprehension monitoring.

The review of the literature also highlighted that older readers demonstrate superior knowledge about reading goals and strategies, and are more able to utilise these skills. They place a greater focus on extracting meaning from text and constructing an integrated representation of the text. In addition, older readers are more likely to flexibly utilise different standards of evaluation to monitor the sense of information. These areas were related to findings in the experiments included in this thesis.

A role for deficient working memory in comprehension failure has been proposed in the literature. However, the relationship between working memory and comprehension monitoring is less clear. Successful comprehension monitoring requires both the storage and processing of information, so that an adequate representation of the text can be constructed. Also, reduced working memory capacity may restrict the amount of integrative processing that can be carried out on discourse. This type of processing is particularly important for the successful monitoring of internal consistency. The relationship between comprehension monitoring and working memory was therefore explored in this thesis.

Finally, work was reviewed which suggests that off-line and real-time measures of assessment may tap different aspects of comprehension monitoring and demand different skills from participants. Perhaps this is unsurprising given the multi-dimensional nature of comprehension monitoring. In comparison to off-line measures, real-time measures may provide a more sensitive means of measuring emerging monitoring skills because they are collected during reading, so measure behaviour that is concurrent with ongoing monitoring processes. Work included in this thesis utilised both off-line and real-time measures, so that a more in-depth insight into comprehension monitoring behaviours could be obtained.

The work presented in the following chapters was designed to gain a better insight into the development of comprehension monitoring in children aged 7 to 10 years, and explore the critical task and reader characteristics mentioned above that may influence this skill.

CHAPTER THREE

OFF-LINE COMPREHENSION MONITORING AND THE RELATION BETWEEN MONITORING ABILITY, READING AND WORKING MEMORY

The first experiment reported in this chapter investigated developmental differences in the detection and identification of different types of error inserted into short texts: nonwords, general knowledge violations and internal inconsistencies. Children were aged 7- to 8-years-old and 9- to 10-years-old. The second experiment explored the relationship between comprehension monitoring ability and independent measures of word reading, reading comprehension and working memory.

3.1. Experiment 1: Off-line comprehension monitoring of nonword, general knowledge violation and internal inconsistency errors

Experiment 1 investigated 7- to 8-year-old and 9- to 10-year-old children's comprehension monitoring skills. Two aspects of comprehension monitoring were measured: Error detection, that is judging whether or not a text contains an error, indicated by responses to questions that evaluate the sense of a text, and error identification, that is underlining errors within a text. The error types were nonwords, general knowledge violations and internal inconsistencies. A paper and pen task was used, described in detail below. This experiment builds upon previous research in a number of ways. First, comprehension monitoring will be investigated developmentally between children representative of the period of transition between beginning and fluent reading. Second, comprehension monitoring will be explored at three different levels: nonword, general knowledge violation and internal inconsistency. Third, the experimental materials in the current experiment address some of the limitations of previous research: Passages were narratives, constructed in a controlled manner that ensured structural consistency across passages, only one error type was inserted into passages, and experimental power was considered through the inclusion of more than one passage for each passage type.

In line with previous research (e.g., Baker, 1984a), it was predicted that in comparison to older children the comprehension monitoring of younger children would be less accurate both in terms of correct responses to a sense question and correct identification of errors. Given the differing ease with which children can adopt the standards of evaluation required for detecting and identifying errors, it was further predicted that comprehension monitoring would be most accurate for passages that contained nonword errors and least accurate for passages that contained internal inconsistency errors. A parallel pattern of performance was expected for responses to sense questions and identifications of errors.

Method

Participants

Ninety-five children (52 boys, 43 girls) with a mean age of 8;0 years (*SD* = 4 months, range = 7;6 – 8;6 years) and 89 children (46 boys, 43 girls) with a mean age of 10;0 years (*SD* = 4 months, range = 9;6 – 10;6 years) participated. The children were enrolled in either Year 3 (7- to 8-years-old) or Year 5 (9- to 10-years-old) in primary or junior schools in England. Children who had special educational needs and those who spoke English as a second language were not included in the experiment. For all participating children informed consent was obtained from school headteachers who acted *in loco parentis* (parents were provided with the opportunity to opt their child out of the experiment if they wished), and children assented to participate.

Materials

Experimental passages were created through the modification of previous comprehension monitoring materials (Language and Reading Research Consortium, in preparation; Oakhill et al., 2005), and the creation of new passages to extend these material sets. There were a total of 51 passages containing errors: 17 for each of the three error types (nonword, general knowledge violation and internal inconsistency), and 24 consistent passages that did not contain any errors (see Appendix 1.1 for a full list of passages). All passages were of a similar length (range = 45 – 55, see Table 3.1) and consisted of five sentences. Passages were narratives in which the opening sentence introduced the topic of the passage, and the last sentence brought the passage to a conclusion. When a narrative involved two characters, a male and a female character were included to eliminate confusion that might arise from the attribution of third person singular pronouns, such as *he* and *she*, to the correct character.

Passage readability. Passages were assessed for readability using the Flesch-Kincaid Grade Level formula (Kincaid, Fishburne, Rogers & Chissom, 1975). Despite criticisms of this formula for focusing on surface features (i.e., number of syllables, words and sentences) of text (e.g., McNamara, Kintsch, Butler-Songer & Kintsch, 1996), it is appropriate for the purposes of the current experiment because alternative measures such as Coh-Metrix (McNamara, Louwerse & Graesser, 2002) that take into account the coherence and comprehensibility of a text would not be suitable for the general knowledge violation and internal inconsistency errors. When preparing passages to be entered into the formula, nonwords were replaced with the word 'thing'. 'Thing' was selected as the replacement word because it allowed all passages to retain a sensible meaning.

The mean Flesch-Kincaid Grade Level of passages for each passage type is shown in Table 3.1. The Grade Level scores were considered appropriate for the reading level of the youngest participants (i.e., 7- to 8-year-old children). A one-way ANOVA with Flesch-Kincaid

Grade Level scores as the dependent variable indicated that there were no significant differences between the Grade Level of the different passage types (nonword, general knowledge violation, internal inconsistency and consistent), $F(3,71) = 1.17$, $p = .33$, $\eta p^2 = .05$. A pilot study with ten adults confirmed that all errors were identifiable.

Table 3.1. Mean numbers (and standard deviations) of words and Flesch-Kincaid Grade Level for nonword, general knowledge violation, internal inconsistency and consistent passages

Passage	**Mean number of words**	**Mean Flesch-Kincaid Grade Level**
Nonword	47.94 (2.28)	2.63 (.31)
General knowledge violation	48.12 (2.85)	2.61 (.28)
Internal inconsistency	49.29 (3.39)	2.56 (.31)
Consistent	48.58 (2.50)	2.72 (.25)

Nonword passages. Seventeen passages containing nonword errors were created by replacing the first, middle or last two consonants of a word within a passage. Replacement letters were the second consonant in the alphabet that followed the consonant being replaced, unless this resulted in a real word, then the process was repeated. For the 17 passages, six passages had the first two consonants, five passages had the middle two consonants, and six passages had the last two consonants of a word replaced. Also, six passages had a nonword in sentence two, six in sentence three and five in sentence four. An example of a nonword passage is shown in Table 3.2. For each passage type (nonword, general knowledge violation, internal inconsistency and consistent), Cronbach's alpha was calculated[1] based on accuracy to the 'did the story make sense' question. For nonword passages Cronbach's alpha was .88, which was acceptable.

[1] Since the passages were experimental, their reliability was unknown. Breakwell, Hammond, Fife-Schaw and Smith (2006) state that reliabilities > .70 are required for a test to be used as a research tool. The reliabilities of the different passage types used in this work exceeded this criterion.

Table 3.2. Example of a passage that contains a nonword

Yesterday Mike played a tennis match.
He hit every danl with great skill, so won the match.
He was given a gold trophy for doing so well.
Mum was very proud and she bought Mike an ice cream.
Mike is looking forward to his next tennis match.

General knowledge violation passages. Seventeen passages containing general knowledge violation errors were created by including information in a sentence that violated common world knowledge. For the 17 passages, four passages had a general knowledge violation in sentence two, seven in sentence three, and six in sentence four. An example of a general knowledge violation passage is shown in Table 3.3. The Cronbach's alpha for general knowledge violation passages was .84, which was acceptable.

Table 3.3. Example of a passage that contains a general knowledge violation

Every morning Jack chops wood for his family.
He always uses a knife to chop the wood.
Jack has to chop at least three logs so that there is enough wood for the fire to burn all day.
On school days Jack has to chop the wood quickly.
Otherwise he will be late for school.

For each general knowledge violation passage, a yes/no question to check that children had the general knowledge related to each error was created (see Appendix 1.2 for a full list of questions). Fifty-eight children (26 7- to 8-year-olds; 32 9- to 10-year-olds) who did not participate in this experiment were asked the background knowledge check questions. Accuracy for all question across the two age groups was considered acceptable (average 96%), and suggested that children in the target age range had the relevant general knowledge to identify the inserted errors.

Internal inconsistency passages. Seventeen passages containing internal inconsistency errors were created by including information in one sentence that contradicted information presented in a previous sentence. Internal inconsistencies always involved information contained in sentences two and four, and one sentence containing information neutral to the inconsistency intervened between these two sentences. An example of an internal inconsistency passage is shown in Table 3.4. The Cronbach's alpha for internal inconsistency passages was .89, which was acceptable.

Table 3.4. Example of a passage that contains an internal inconsistency

Jack has a grey rabbit named Bugs.
<u>Bugs lives in his cage in the house and never goes outside.</u>
Jack feeds Bugs rabbit food and a carrot every Friday for a treat.
<u>Everyday Bugs plays in the garden on the grass.</u>
Jack really likes Bugs.

Consistent passages. Twenty-four consistent passages were created that contained no errors. An example of a consistent passage is shown in Table 3.5. The Cronbach's alpha for consistent passages was .83, which was acceptable.

Table 3.5. Example of a consistent passage

Charlie likes tropical fish.
In his bedroom, he has three fish that live in a big glass tank.
Every morning before school Charlie feeds the fish.
Once a week Charlie helps Mum to clean the tank.
Charlie likes to watch the pretty colours of the fish as they swim.

Procedure

Passages were assembled into three different booklets, each containing 24 experimental passages: 16 error passages (with either five or six items for each of the three error types, see Appendix 1.3 for further details) and eight consistent passages. All passage types were included in each booklet so that children's performance could be compared across the different error types. The order in which passages were presented across the three booklets was the same for all but the last passage, where a different error type was presented in each booklet. In each booklet the first three passages were arranged in a constrained order such that a nonword passage was encountered first, a general knowledge passage second and an internal inconsistency passage third. The remaining passages were presented in a pseudorandomised order in which no more than three passages of the same type occurred consecutively within a booklet. A one-way ANOVA with Flesch-Kincaid Grade Level scores as the dependent variable indicated that there were no significant differences between the Grade Level of passages for the three booklets, $F(2,69) < 1$, ns.

Each page of a booklet was headed by the word 'story' followed by the number 1-24, and displayed a single passage. Text was presented in black font on a white background. Below each passage was the question 'did the story make sense?' and two cartoon faces, one with a happy and one with a sad expression. These faces were used by children to indicate whether or not a passage made sense. Rather than using a continuum of facial expressions, a happy and a sad expression were selected so that children would focus on a good or bad sense threshold.

The passage booklets were administered to children in year group classes in their school classroom. The whole session lasted approximately 35 minutes. Each child completed only one booklet. Booklets were allocated as evenly as possible among participants. Children were instructed to read the passages silently in their heads and underline any 'silly mistakes'. Such 'silly mistakes' were considered to be parts of the text (i.e., words or sentences) that children did not understand or that did not make sense. Before beginning the task, the experimenter talked the children through three practice passages, one for each error type (nonword, general knowledge violation and internal inconsistency). Children who had difficulty with the practice passages put up their hands and received guidance from the experimenter who reiterated the task of underlining 'silly mistakes' in the passages until the child understood what they were required to do. After reading each passage children responded to the question 'did the story make sense?' by ticking a happy or sad face. Children were told that they should tick the happy face if the passage did not contain a 'silly mistake' and tick the sad face if the passage contained a 'silly mistake'.

Scoring

The three error types (nonword, general knowledge violation and internal inconsistency) were scored with either 1 for the correct underlining of the error or 0 for failing to underline the error. For all passage types (nonword, general knowledge violation, internal inconsistency and consistent) incorrect error identifications were scored with 1 for each part of a text (i.e., word or phrase) that was incorrectly underlined. Approximately 3% of children incorrectly underlined more than two parts of a text. A score of 1 was awarded for correctly ticking a sad or a happy face and a score of 0 for incorrectly ticking a sad or a happy face.

Results

Before analysis the Q-Q plots for all measures were examined for each age group separately. Q-Q plots provide a graphical means of exploring the distribution of a data set created by plotting quantiles of a variable against the quantiles of a normal distribution (e.g., Cohen, Cohen, West & Aiken, 2003). If values fall on the diagonal of the plot and form a line that is roughly straight, then the variable has a normal distribution. Whereas, if the values deviate substantially from the diagonal of the plot problems with kurtosis are indicated when values form a line that consistently sags below or above the diagonal, and skewness are indicated when values form an S-shaped curve (e.g., Field, 2005). Based on this criteria, the plots indicated that the distribution of the data was acceptable.

Children did not read the same number of passages for each passage type[2]. Thus, for ease of comparison, the correct sense judgement, correct error identification and incorrect error identification data was converted into proportions (number of correct sense judgements, correct error identifications and incorrect error identifications, respectively, as a proportion of the total number of passages read for each passage type). The analyses that follow were performed on both untransformed and transformed (arcsine because of proportional data) data sets. For all analyses, the pattern of results was the same for both untransformed and transformed data sets. Therefore, throughout this Results Section the untransformed results are reported. Note, here and in all subsequent analyses within this thesis main effects involving more than two conditions were analysed by *t*-tests to identify significant differences between conditions. Where appropriate, multiple *t*-tests were corrected for family-wise error rate by adjusting each *p*-value (i.e., the .05 *p*-value was divided by the number of tests).

---

[2] Note, this was due to there being different numbers of error and consistent passages, and each of the booklets had a different number of passages for one of the error types.

Correct sense judgements

This analysis focuses on responses to the question 'did the story make sense?'. The required response was the sad face for error passages, and the happy face for consistent passages. The mean proportions of correct sense judgements made for each passage type by age group are shown in Table 3.6. The mean proportion of correct sense judgements for nonword passages for the 9- to 10-year-old children was close to ceiling, so this passage type was not included in the analysis. The mean correct for the general knowledge violation, internal inconsistency and consistent passage types was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA.

Table 3.6. Mean proportions (with standard deviations) of correct judgements by age group and passage type

	Age group		
Passage	**7- to 8-year-olds (n = 95)**	**9- to 10-year-olds (n = 89)**	**Total**
Nonword	.86 (.20)	.94 (.13)	.90 (.17)
General knowledge violation	.84 (.19)	.88 (.20)	.86 (.20)
Internal inconsistency	.71 (.26)	.84 (.23)	.78 (.25)
Consistent	.81 (.18)	.90 (.14)	.86 (.16)
Total	.81 (.21)	.89 (.18)	

There was a significant main effect of age group ($F(1,182) = 22.33$, $p < .001$, $\eta p^2 = .11$) because the older children made more correct sense judgements than the younger children. There was also a main effect of passage type ($F(2,364) = 10.93$, $p < .001$, $\eta p^2 = .06$). Children made a similar proportion of correct sense judgements for the consistent and general knowledge violation passages ($t(183) < 1$, ns), and there were more correct sense judgements for these two passage types than for the internal inconsistency passages ($t(183) = 3.60$, $p < .001$, $d = .38$, for consistent and internal inconsistency passages; $t(183) = 4.41$, $p < .001$, $d = .35$ for general knowledge violation and internal inconsistency passages). The interaction between age group and passage type was not significant ($F(2,364) = 2.16$, $p = .12$, $\eta p^2 = .01$).

Correct error identifications

This analysis focuses on correct identification of the deliberately inserted errors, so scores only for the nonword, general knowledge violation and internal inconsistency passages are considered. The mean proportions of correct error identifications for each passage type by age group are shown in Table 3.7. Again, the mean proportion of correct identifications for nonword errors for the 9- to 10-year-old children was close to ceiling, so this passage type was not included in the analysis. The mean correct for the general knowledge violation and internal inconsistency passage types was the dependent variable in a 2 (age group) x 2 (passage type) ANOVA.

There was a significant main effect of age group ($F(1,182) = 24.04$, $p < .001$, $\eta p^2 = .12$) because older children made more correct identifications than younger children. However, the main effect of passage type did not reach significance ($F(1,182) = 3.21$, $p = .08$, $\eta p^2 = .02$). The effect of age group was qualified by a significant interaction with passage type ($F(1,182) = 10.62$, $p = .01$, $\eta p^2 = .06$), which is shown in Figure 3.1. The interaction arose because older children showed no effect of passage type ($t(88) = 1.16$, $p = .25$, $d = .13$), whereas younger children showed an effect of passage type ($t(94) = 3.31$, $p = .01$, $d = .47$).

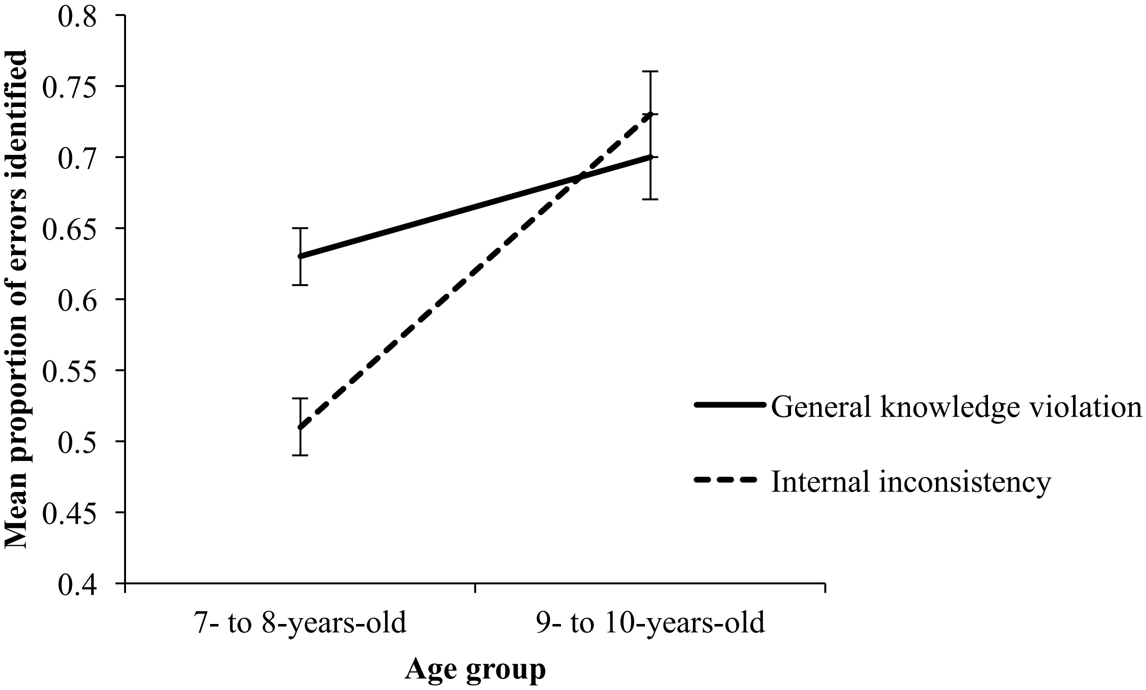

Figure 3.1. Mean proportion (+/- standard error) of errors correctly identified by age group and passage type

The relation between correct sense judgements and correct error identifications

To explore the relationship between correct sense judgements and correct error identifications correlations were calculated for each age group separately. For both age groups, the ability to judge that a passage contained an error was strongly correlated with the identification of nonword (*rs* = .88 and .79, *ps* < .001, for older and younger age groups, respectively), general knowledge violation (*rs* = .78 and .65, *ps* < .001) and internal inconsistency (*rs* = .80 and .66, *ps* < .001) errors.

Table 3.7. Mean proportions (with standard deviations) of correct identifications and incorrect identifications by age group and passage type

	Correct identifications			Incorrect identifications		
Passage type	**7- to 8-year-olds (n = 95)**	**9- to 10-year-olds (n = 89)**	**Total**	**7- to 8-year-olds (n = 95)**	**9- to 10-year-olds (n = 89)**	**Total**
Nonword	.82 (.25)	.94 (.14)	.88 (.20)	.25 (.30)	.10 (.18)	.18 (.24)
General knowledge violation	.63 (.22)	.70 (.21)	.67 (.22)	.17 (.24)	.06 (.10)	.12 (.17)
Internal inconsistency	.51 (.30)	.73 (.26)	.62 (.28)	.24 (.28)	.10 (.14)	.17 (.21)
Consistent	-	-	-	.23 (.27)	.12 (.19)	.18 (.23)
Total	.65 (.26)	.79 (.20)		.22 (.27)	.10 (.15)	

Note. Cells are empty for correct identifications for consistent passages because these passages did not contain errors to be identified.

Incorrect error identifications

There were instances when children underlined a word or phrase that was not the target error. Such incorrect error identifications could occur in the passages with inserted errors and also in the consistent passages with no errors. The mean proportions of incorrect error identifications for each passage type by age group are shown in Table 3.7. In both age groups there were few instances when children made incorrect error identifications. The mean proportion of incorrect identifications for the four passage types was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA.

There was a significant main effect of age group ($F(1,182) = 28.95$, $p < .001$, $\eta p^2 = .14$) because the older children made fewer incorrect identifications than the younger children. There was also a main effect of passage type ($F(3,546) = 5.26$, $p = .01$, $\eta p^2 = .03$). Children made a similar number of incorrect identifications within consistent, nonword and internal inconsistency passages (all $ts(183) < 1$, ns), and fewer incorrect identifications within general knowledge violation passages (all $ts(183) > 3.18$, $ps < .01$). The interaction between age group and passage type was not significant ($F(3,546) < 1$, ns).

Discussion

The main aim of Experiment 1 was to determine whether developmental differences exist in comprehension monitoring, measured by children's ability to detect and identify different types of error in text. As predicted, developmental differences were apparent: Older children made significantly more correct sense judgements and correctly identified significantly more errors (for all error types) than younger children. It should be noted that young children made more incorrect error identifications than older children. Also in line with predictions, children made more correct sense judgements for nonword passages and fewer correct sense judgements for internal inconsistency passages. Children made a similar proportion of correct sense judgements for consistent and general knowledge violation passages. Further, children made a similar proportion of correct error identifications for general knowledge violation and internal inconsistency passages. The ability to correctly judge that a passage contained an error was significantly correlated with error identification for all error types. With respect to incorrect identifications, children made a similar proportion of incorrect identifications for consistent, nonword and internal inconsistency passages, and fewer incorrect identifications for general knowledge violation passages.

In line with previous research (e.g., Baker, 1984a, 1984b), findings revealed developmental differences in children's comprehension monitoring. Older children made significantly more correct sense judgements, and correctly identified significantly more errors (for all error types) than younger children. Such developmental differences may be accounted for by younger children's developing information processing and metacognitive skills. In comparison to older children, younger children's comprehension monitoring abilities may be restricted by information processing limitations, a reduced knowledge of reading comprehension processes and a lack of insight into how errors impair their understanding (e.g., Perfetti et al., 1996).

In addition, the finding that younger children made more incorrect error identifications than older children is of interest because it suggests that younger children's inability to identify errors was not a result of their reluctance to point out and underline the problems within the passages. Younger children attempted to undertake the task, however their accuracy in identifying errors was lower than that of the older children. These results are consistent with previous research using off-line measures (e.g., Baker, 1984a) in showing that younger children do not engage in accurate monitoring as frequently as older children, and thus are poorer at monitoring their comprehension of texts.

As predicted, both age groups were better at detecting and identifying nonword errors and least accurate at detecting and identifying internal inconsistency errors. For younger children, there was a significant difference between the identification of general knowledge violation and internal inconsistency errors. In line with previous research (e.g., Baker, 1984a), younger children identified more general knowledge violation than internal inconsistency errors. However, this was not the case for older children who identified a comparable number of general knowledge violation and internal inconsistency errors. It seems that the source of this age group difference was better identification of internal inconsistencies by older children.

There are several possible explanations for the differences in monitoring performance across the error types (nonword, general knowledge violation and internal inconsistency), and the age group differences in monitoring performance on the general knowledge violation and internal inconsistency errors. One explanation relates to the standards of evaluation children are required to adopt when detecting and identifying different error types, and the different information processing demands associated with each standard. As mentioned in the Literature Review, adopting a lexical standard is less demanding of information processing resources than adopting an external consistency standard, which is less demanding of information processing resources than adopting an internal consistency standard. Due to children's developing processing capacity, young children may have insufficient resources to consistently and effectively adopt more demanding standards of evaluation. Thus, in comparison to nonword and general knowledge violation errors children may struggle to monitor their comprehension for internal inconsistency errors, and this may be particularly true for younger children whose information processing capabilities are less developed.

Children's reading goals may also impact upon the monitoring of different error types. Young children's reading goals tend to focus on the decoding of words and the meaningfulness of individual sentences, rather than undertaking monitoring behaviours such as integrating information within and across sentences (e.g., Myers & Paris, 1978). Adopting such goals supports the successful monitoring of nonword and general knowledge violation errors, but fails to foster the monitoring of internal inconsistency errors. Young children's failure to adopt adequate reading goals may be due to their developing metacognitive skills, which limits their knowledge of reading comprehension processes. Thus, young children may lack the knowledge or ability to independently adopt more complex reading goals, or simultaneously utilise a range of different reading goals necessary for the detection and identification of nonword, general knowledge violation and internal inconsistency errors. However, information processing capacity limitations may also influence children's reading goals. Young children may have insufficient cognitive resources to enable them to adopt more demanding reading goals, which may result in monitoring problems when encountering more complex error types.

In addition, another possible explanation for poorer monitoring of general knowledge violation and internal inconsistency errors is that children may detect the problem, but make an assumption to resolve it. Younger children in particular may accept the error information as being true so that it no longer violates general knowledge, or embellish the error information so that it is no longer inconsistent. Consequently, children may fail to detect and identify general knowledge violation or internal inconsistency errors because they no longer perceive error information as problematic (e.g., Markman, 1979).

Furthermore, the relationship between correct sense judgements and correct error identifications implies that children who correctly judged that a passage contained an error were likely to correctly identify the error. However, for both age groups, there was a difference between the number of correct sense judgements and the number of correct error identifications. Children made more correct sense judgements than correct error identifications. This finding suggests that children claim to detect errors in texts, but do not always accurately identify them. As proposed by Oakhill et al. (2005), one possible reason for this pattern is that children only had a vague sense of something being wrong with the passages (i.e., a feeling of confusion, the reason for which they were unable to pinpoint). Such a reason is supported by Ruffman's (1996) suggestion that children correctly judge a text to be problematic, but fail to identify the error because they lack the metacognitive insight to understand why the text does not make sense. In addition, despite every effort being made to ensure that children undertook the task with care and attention, as the experiment was conducted in a group setting in the children's classroom, it may be that children rushed the task to complete it as soon as possible. Thus, they may not have paid full attention to underlining the correct information.

This was the first experiment that included the experimental passages, so there was a focus on developing and testing the passages, particularly the error passages. Thus, this experiment included more error than consistent passages. It should be noted that presenting children with different numbers of error and consistent passages may bias their reading or comprehension monitoring strategies. One possibility is that children may become aware that passages contain errors and undertake a more purposeful reading of the text. However, children received instructions alerting them to the possibility that passages may contain 'silly mistakes', so it may be argued that children were already provided with information that encouraged them to adopt a purposeful reading of the text that included monitoring their comprehension. Another possibility is that children may show a response bias to the 'did the story make sense' question, responding with 'no' rather than 'yes'. If such a bias was present in this experiment, correct responses to the sense question for all of the error passages (nonword, general knowledge violation and internal inconsistency) would be at ceiling. This was not the case; analysis of the correct sense judgements revealed that only the correct sense judgements for nonword passages for the 9- to 10-year-old children were close to ceiling. Rather than this finding resulting from a response bias, it is possible that older children have mastered the ability to adopt a lexical standard and monitor their comprehension for nonword errors.

The results of Experiment 1 demonstrate that there are developmental differences in children's comprehension monitoring ability. Younger children do not engage in accurate monitoring as frequently as older children, and thus are poorer at detecting and identifying errors within a text. There are also effects of error type; children were able to proficiently monitor their comprehension for nonword errors, but were poorer, especially in the case of younger children, at monitoring general knowledge violation and internal inconsistency errors. It is possible that these age group and error type findings may be due to information processing factors. However, Experiment 1 did not include an independent assessment of information processing capacity, such as working memory. Thus, Experiment 2 was conducted to explore the contribution of working memory, along with word reading and reading comprehension skills, to comprehension monitoring ability.

3.2. Experiment 2: The relation between comprehension monitoring ability, reading and working memory

Investigations of individual differences have identified a wide range of skills that are related to children's reading comprehension level. However, little research has investigated the skills that are related to children's comprehension monitoring ability. Experiment 2 focused on the contribution of word reading, reading comprehension and working memory to children's comprehension monitoring ability. To do this, the performance of the children who participated in Experiment 1 was examined, and the best and worst comprehension monitors (hereafter good and poor comprehension monitors, respectively) were identified on the basis of their correct sense judgement and correct error identification performance. These groups were then compared on independent measures of word reading, reading comprehension and working memory, and the relationship between these skills and monitoring ability was explored. The rationale and aims are detailed below.

There is a strong predictive relationship between children's word reading and reading comprehension level in the early years (e.g., Juel, Griffith & Gough, 1986). For this reason, it is likely that young children's ability to proficiently monitor their comprehension will be influenced by their word reading skill. Slow or inaccurate word reading may affect comprehension monitoring by placing too large a demand on children's information processing capacity, leaving little remaining for undertaking important monitoring processes such as integration (e.g., Perfetti, 1985).

In addition, it is likely that young children's comprehension monitoring proficiency is related to their reading comprehension level. Poor or inaccurate comprehension of a text may reflect impaired comprehension monitoring ability, as it is unlikely that a child can construct an integrated and coherent representation of a texts meaning without engaging in adequate comprehension monitoring (e.g., Cain et al., 2004). However, it may be the case that a child needs to be a good comprehender, able to construct an adequate situation modal, in order to engage in proficient comprehension monitoring. To-date, the direction of the causal relationship between comprehension monitoring skill and the construction of an adequate situation model remains unclear.

Further, the relative importance of word reading and reading comprehension skills for comprehension monitoring proficiency may change during the course of development, and these skills may differ in their associative contribution for good and poor comprehension monitors. Thus, Experiment 2 included independent assessments of word reading and reading comprehension in order to determine whether word reading, both in isolation and in context, and reading comprehension skills make a distinct contribution to comprehension monitoring ability.

As discussed in the Literature Review, it has been suggested that young children's success in comprehension monitoring may be influenced by the working memory demands of detecting and identifying different types of errors (e.g., Perfetti, et al., 1996; Seigneuric et al., 2000). A child's ability to fully monitor their comprehension requires both the storage and processing of information. Previously read information must be held in memory while new incoming information is encoded and integrated into the situation model (e.g., Yuill & Oakhill, 1991). For this reason, children who have poor working memory may be poor at monitoring their comprehension. Consequently, these children may be able to detect and identify errors when the working memory demands of a task are low, for example, nonword and general knowledge violation errors as the error information should be familiar, so it is likely to have a stable and accessible memory representation. However, comprehension monitoring ability may be hindered when working memory demands of a task are high, for example, internal inconsistency errors, as the error information is newly presented and must be compared with previously read information. In contrast, it has been proposed that readers may demonstrate impaired comprehension monitoring despite possessing adequate working memory resources (e.g., Perfetti et al., 1996). Thus, some children may simply have a fundamental problem with comprehension monitoring in the presence of adequate memory.

A handful of studies investigating children's comprehension monitoring have included an independent assessment of working memory (e.g., Oakhill et al., 2005; van der Schoot et al., 2011). However, in the main these studies have only included one age group and focused on good and poor comprehenders. Findings have shown that working memory plays a role in differentiating group performance on comprehension monitoring tasks. Other research (e.g., Cain et al., 2004) has shown that for reading comprehension in general, working memory and comprehension monitoring share unique variance and also predict unique variance. This suggests that working memory may not explain variance in comprehension monitoring tasks, rather other factors such as metacognitive skills may be playing a role. To further investigate the relationship between comprehension monitoring and working memory, it was informative to explore working memory in two different age groups that vary in monitoring ability.

The following predictions were made. In line with previous research (e.g., Oakhill et al., 2005), word reading, reading comprehension and working memory were predicted to differ between the groups of good and poor comprehension monitors in both age groups. Further, if the relationship between comprehension monitoring and working memory is due to a linguistic processing component, it was predicted that groups would be differentiated more strongly on the verbal than the numerical working memory measure.

Method

Participants

Children from Experiment 1 identified as having the very best (good comprehension monitors) and the poorest (poor comprehension monitors) comprehension monitoring participated. To categorise the groups, the mean proportions of correct error identifications for each error type (nonword, general knowledge violation and internal inconsistency) were converted to Z-scores. Good comprehension monitors were considered to be 7- to 8-year-old children with positive Z-scores (i.e., original scores above the mean) for nonword, general knowledge violation and internal inconsistency errors, and 9- to 10-year-old children with positive Z-scores for general knowledge violation and internal inconsistency errors. Poor comprehension monitors were considered to be 7- to 8-year-old children with negative Z-scores (i.e., original scores below the mean) for nonword, general knowledge violation and internal inconsistency errors, and 9- to 10-year-old children with negative Z-scores for general knowledge violation and internal inconsistency errors. Due to the amount of time it is reasonable to work within one school, 15 children with the highest positive Z-scores and 15 children with the lowest negative Z-scores for each error type in each age group were selected as good and poor comprehension monitors, respectively (7- to 8-year-olds: Good: 9 boys, 6 girls, mean age 8;0 years, *SD* = 3 months, range = 7;6 – 8;5 years; Poor: 7 boys, 8 girls, mean age 8;0 years, *SD* = 3 months, range = 7;7 – 8;5 years, and 9- to 10-year-olds: Good: 6 boys, 9 girls, mean age 10;2 years, *SD* = 3 months, range = 9;1 – 10;6 years; Poor: 9 boys, 6 girls, mean age 10;0 years, *SD* = 4 months, range = 9;6 – 10;6 years). Informed parental consent was obtained for each child, and children assented to participate.

To confirm that good and poor comprehension monitors in each age group differed in monitoring behaviour, independent samples *t*-tests were conducted for correct sense judgements and error identifications for each error type. A summary of these statistics, along with the mean proportions of correct sense judgements and error identifications made for each error type by good and poor comprehension monitors in each age group, are shown in Table 3.8. Good and poor comprehension monitors aged 7 to 8 years differed significantly on the proportions of correct sense judgements and identifications of nonword errors. Differences on nonword performance between the good and poor comprehension monitors aged 9 to 10 years were not assessed, as nonword errors were not used to identify monitor ability in this age group. Good and poor comprehension monitors, in both age groups, differed significantly on the proportions of correct sense judgements and identifications of general knowledge violation and internal inconsistency errors.

Table 3.8. Mean proportions (with standard deviations and t-tests) of correct judgements and identifications by age group, comprehension monitor group and error type

		7- to 8- year-olds				**9- to 10-year-olds**			
		Good	**Poor**			**Good**	**Poor**		
Error type	**Monitoring measure**	**(n = 15)**	**(n = 15)**	***t*(28)**	***d***	**(n = 15)**	**(n = 15)**	***t*(28)**	***d***
Nonword	**Judgements**	.97 (.07)	.63 (.22)	5.67**	2.08	.99 (.05)	.96 (.08)	-	-
	Identifications	.98 (.06)	.56 (.23)	6.89**	2.50	1.00 (.00)	.94 (.16)	-	-
General knowledge violation	**Judgements**	.98 (.06)	.74 (.24)	3.72*	1.37	.94 (.10)	.73 (.20)	3.70*	1.33
	Identifications	.85 (.12)	.43 (.15)	8.46**	3.09	.90 (.08)	.46 (.14)	10.41**	3.86
Internal inconsistency	**Judgements**	.88 (.13)	.59 (.33)	3.19*	1.16	.97 (.07)	.58 (.32)	4.62**	1.68
	Identifications	.80 (.14)	.38 (.23)	6.04**	2.21	.96 (.07)	.46 (.23)	8.04**	2.94

Note. In the older age group, results for the *t*-tests between good and poor comprehension monitors correct judgements and identifications of nonword errors are not reported because these differences were not assessed (see Participant Section for rationale). $^{*}p < .01$, $^{**}p < .001$.

Procedure

Children were tested individually at school in a quiet room, separate from their classroom. During two 15 minute sessions, children completed measures of word reading, reading comprehension and working memory described below. All measures were scored in accordance with manual guidelines.

Word reading (in isolation) measures. To assess the accuracy and fluency of word reading in isolation the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner & Rashotte, 1999) Form A was administered. In the Sight Word Efficiency subtest children were given 45 seconds to read aloud as many real words as possible, printed in a series of four vertical lists. In the Phonemic Decoding Efficiency subtest children were given 45 seconds to read aloud as many pronounceable nonwords as possible, printed in a series of three vertical lists. For both tests, if a child paused on a word for three seconds they were prompted by the experimenter to move on to the next word. The test-retest reliability coefficients for Form A subtests reported in the manual for the age groups of the children in the current experiment range from .84 to .97. For the word reading in isolation measures the dependent variable was the number of words or nonwords that were read correctly.

Word reading (in context) and reading comprehension measures. The York Assessment of Reading Comprehension Passage Reading (YARC Passage Reading; Snowling et al., 2009) was administered to measure reading rate (i.e., the time taken to read a passage, in seconds), reading accuracy (i.e., the number of word reading errors) and reading comprehension (i.e., the number of correct responses to comprehension questions). Note that the reading rate measure will be presented in descriptive tables within this thesis, but it will not be used in analysis. This is because the experimenter is able to provide guidance when children make reading errors or are unable to read words, and children adopt different oral reading strategies in response to this guidance, which has differing effects on their reading rate.

It was possible that some of the children identified as having poor comprehension monitoring would have reading comprehension problems. In such cases, the YARC Passage Reading manual suggests that a more reliable estimate of children's ability can be obtained by assessing comprehension on two parallel passages of equal difficulty, rather than two consecutive passages of differing difficulty. Thus, each age group read two passages at a level appropriate for their school year group (i.e., 7- to 8-year-olds read two Level 3 passages and 9- to 10-year-olds read two Level 5 passages). After reading each passage, children responded to eight comprehension questions.

The reliability coefficients for the passages used range from .92 to .94 for word reading rate and reading accuracy, and .48 to .62 for reading comprehension. Although the values for the reading comprehension measure are a little low, this is perhaps not surprising given the range of questions included in the YARC Passage Reading. Questions tap inferential skills, literal memory and vocabulary knowledge. As a result, these values are most probably an indication of the multi-dimensional nature of text comprehension. For analysis three scores were calculated: reading rate, reading accuracy and reading comprehension.

Working memory measures

Verbal working memory measure. The Sentence Span task from the Working Memory Test Battery for Children (WMTB-C; Pickering & Gathercole, 2001) was administered. Children were asked to listen to lists of unrelated sentences (3 – 5 words in length). Immediately after hearing each sentence, children had to state whether it was true or false, e.g., books have pages (true); fish have long hair (false) (the processing component). When all of the sentences in a trial had been heard, children had to recall the final word from each sentence, in the correct serial order (the storage component). The task began with two practice trials each at a difficulty level of one sentence and two sentences. If a child correctly recalled the final words (in their correct serial order) at the two-sentence difficulty level, they started the test items at the two-sentence level; if not they started at the one-sentence level. Appropriate feedback was given on the practice trials, but no feedback was given during the test proper. Difficulty increased to a maximum level of six sentences in a trial. There were six test trials for each level of task difficulty. The experimenter moved on to the next level of difficulty when three trials at a given level were correctly recalled. Testing was discontinued when children incorrectly recalled four trials within one level of difficulty. The mean test-retest reliability coefficient reported in the manual is .61.

Numerical working memory measure. The Backward Digit Span task from the WMTB-C was administered. Children listened to a list of digits from one to nine and were asked to recall the digits in reverse order. The task began with two practice trials each at a difficulty level of two and three digits. If a child correctly recalled the digits (in their correct serial order) at the three-digit difficulty level, they started the test items at the three-digit level; if not they started with two-digit level. Appropriate feedback was given on the practice trials, but no feedback was given during the test proper. Task difficulty increased to a maximum level of seven digits in a string. There were six trials at each level of difficulty. The experimenter moved on to the next level of difficulty when three trials at a given level were correctly recalled. Testing was discontinued when children incorrectly recalled four trials within one level of difficulty. The mean test-retest reliability coefficient reported in the manual is .81.

For both of the working memory measures there were three dependent variables: The number of items that were correctly recalled, the number of trials that were correctly recalled and the maximum level of difficulty (i.e., span) reached.

Results

Before analysis the Q-Q plots for all measures were examined for each age group separately. The plots indicated that the distribution of the data was acceptable. Table 3.9 and 3.10 summarise the mean performance of the good and poor comprehension monitors (by age) on the word reading and reading comprehension, and memory measures, respectively. After collecting the data, a revised manual for the YARC Passage Reading was published which omits the advice to administer two parallel passages and does not support the calculation of standard scores for reading comprehension for this administration. It was therefore not appropriate to calculate standard scores for the YARC Passage Reading comprehension measure because of the procedure followed and the revised manual guidelines[3]. Because groups are tightly matched for age (7- to 8-year-old good and poor comprehension monitors mean age = 8;0 years; 9- to 10-year-old good and poor comprehension monitors mean age = 10;2 and 10;0 years, respectively) and children in each age group completed the same passage level of the YARC, all analyses were conducted on raw scores. Independent samples *t*-tests were conducted to test for differences in the word reading, reading comprehension and memory measures for each age group, separately.

[3] In light of this, the YARC Passage Reading was correlated with an alternative assessment of children's reading comprehension ability: the Neale Assessment of Reading Ability (NARA; Neale, 1999). Twenty-six children (14 7- to 8-year-olds; 12 9- to 10-year-olds) who did not participate in this experiment completed both assessments. Strong correlations between the standard scores for the reading rate, reading accuracy and reading comprehension measures were found (*rs* = .83, .76 and .61, *ps* < .01, for reading rate, reading accuracy and reading comprehension measures, respectively). The strength of the correlations between the two different assessments indicates that the YARC Passage Reading is a valid measure of reading comprehension.

Table 3.9. Mean raw scores, ability scores and standardised scores (with standard deviations) for word reading and reading comprehension measures by age group and monitor ability

	7- to 8-year-olds		9- to 10-year-olds	
Measure	**Good** (n = 15)	**Poor** (n = 15)	**Good** (n = 15)	**Poor** (n = 15)
Word reading				
Raw score (maximum = 104)	65.20 (6.35)	57.40 (14.25)	67.53 (7.87)	66.40 (6.61)
Standardised score	117.33 (6.87)	110.07 (11.58)	102.20 (7.66)	103.13 (10.62)
Nonword reading				
Raw score (maximum = 63)	34.27 (9.15)	27.60 (11.27)	40.60 (7.32)	35.13 (8.04)
Standardised score	114.93 (10.30)	106.93 (11.97)	110.20 (8.13)	104.40 (11.81)
Reading rate				
Raw score	80.53 (22.02)	115.73 (55.40)	106.53 (32.96)	134.90 (34.79)
Ability score	72.60 (7.40)	64.07 (14.80)	80.87 (8.20)	73.00 (7.18)
Standardised score	117.63 (7.27)	108.70 (12.34)	111.90 (10.61)	101.80 (10.59)
Word reading accuracy				
Raw score	3.73 (2.82)	8.57 (5.15)	4.63 (4.15)	7.97 (5.14)
Ability score	53.13 (6.12)	46.07 (8.61)	63.83 (7.59)	58.63 (8.84)
Standardised score	109.20 (7.80)	100.23 (10.57)	109.23 (9.86)	101.67 (11.69)
Reading comprehension				
Raw score (maximum = 8)	5.13 (1.32)	4.20 (1.36)	4.03 (1.08)	3.33 (1.16)

Note. Raw scores are presented on the upper row, and where available ability score on the middle row and standardised scores on the lower row.

Word reading (in isolation) measures

The same pattern of performance was found for both age groups. The groups of good and poor comprehension monitors did not differ significantly on the word reading or the nonword reading measures (all $ts(28) < 1.95$, ns). Note that all children had typical word and nonword reading above a standard score of 85.

Word reading (in context) and reading comprehension measures

The performance of the younger good and poor comprehension monitors differed significantly on the reading accuracy measure ($t(28) = 3.19$, $p = .01$, $d = 1.17$). However, the difference in performance between the older good and poor comprehension monitors was only marginally significant on this measure ($t(28) = 1.96$, $p = .06$, $d = .72$). Surprisingly, the difference in performance between the good and poor comprehension monitors in both age groups did not reach conventional levels of significance on the reading comprehension measure (both *ts*(28) < 1.95, *ps* > .07). Note that all children had word reading equal to or above a standard score of 84.

Table 3.10. Mean raw scores and standardised scores (with standard deviations) for working memory measures by age group and monitor ability

	7- to 8-years-old		9- to 10-years-old	
Measure	**Good**	**Poor**	**Good**	**Poor**
Verbal working memory				
No. of items (maximum = 126)	19.36 (8.68)[a]	15.80 (5.66)	27.93 (11.98)	27.00 (7.90)
No. of trials (maximum = 36)	10.64 (3.25)[a]	9.87 (2.10)	13.27 (3.77)	13.47 (2.64)
No. of trials (standardised score)	99.14 (14.75)[a]	95.33 (10.50)	107.47 (20.47)	108.40 (14.28)
Span (maximum = 6)	1.79 (.58)[a]	1.73 (.46)	2.27 (.70)	2.27 (.59)
Numerical working memory				
No. of items (maximum = 169)	48.87 (17.65)	41.00 (15.68)	61.33 (23.76)	58.53 (24.51)
No. of trials (maximum = 36)	14.07 (3.84)	12.00 (3.44)	16.73 (4.83)	15.93 (4.56)
No. of trials (standardised score)	109.67 (15.45)	100.40 (16.29)	113.80 (19.71)	110.33 (18.83)
Span (maximum = 7)	3.53 (.74)	3.13 (.74)	3.80 (.77)	3.87 (.99)

Note. Unless indicated in the table by superscript, n = 15 for both age groups. [a] n = 14. Missing data is due to one child not completing the verbal working memory measure.

Working memory measures

The same pattern of performance was found for both age groups. The groups of good and poor comprehension monitors did not differ significantly on any of the memory measures (all *ts*(28) < 1.55, ns). Note that all children had typical verbal and numerical working memory above a standard score of 85.

Discussion

Experiment 2 explored the relationship between comprehension monitoring ability and independent measures of word reading, reading comprehension and working memory in 7- to 8-year-old and 9- to 10-year-old children identified as good and poor comprehension monitors. Good and poor comprehension monitors in both age groups did not differ significantly on the word reading, nonword reading, reading comprehension or verbal and numerical working memory measures. However, the younger group of good and poor comprehension monitors differed significantly on the reading accuracy measure. Whilst this section discusses these findings, it should be noted that the null differences in this experiment are difficult to interpret given the small number of participants. Thus, it is possible that null differences between good and poor comprehension monitors simply reflect a lack of statistical sensitivity.

Findings provide partial support for the prediction that good and poor comprehension monitors would differ in terms of word reading skill. For the younger age group, there was a significant difference between good and poor comprehension monitors on the reading accuracy measure. However, the difference did not reach conventional levels of statistical significance for the older group. As suggested by Perfetti (1985) accurate word reading skills free up processing resources that can be devoted to monitoring, rather than word decoding. The difference in reading accuracy in the younger age group lends support to this hypothesis.

For the younger age group, the poor comprehension monitors' significantly lower performance on the reading accuracy measure is of interest when compared to the non-significant difference between good and poor comprehension monitors performance on word and nonword reading measures. This pattern of results suggests that in younger children, poor comprehension monitors are able to efficiently and accurately read words in isolation, but struggle when reading words in context. Thus, it would seem that problems arise for younger poor comprehension monitors when trying to read for meaning and construct a rich and coherent situation model.

In contrast to the predictions, there was no significant difference in reading comprehension performance between the good and poor comprehension monitors for either age group. Such a finding is surprising as comprehension monitoring would be expected to be an integral skill in achieving a full and coherent representation of a text necessary for adequate comprehension. Indeed, previous research has found differences in the comprehension monitoring abilities of good and poor comprehenders (e.g., Oakhill et al., 2005; van der Scoot et al., 2011). However, it should be noted that the current experiment selected participants in a different manner to previous research, such as that noted above, as selection was based on comprehension monitoring ability not reading comprehension ability.

Further, in contrast to the predictions, for both age groups there were no significant differences between good and poor comprehension monitors across the verbal and numerical working memory measures. While this may suggest that there are no differences in basic information processing abilities between good and poor comprehension monitors, it is unlikely that this is the case as previous research (e.g., Oakhill et al., 2005; van der Schoot et al., 2011) has documented the importance of verbal and numerical working memory in comprehension monitoring ability. Within the current experiment the working memory demands of the comprehension monitoring task were low. Passages were five sentences in length, and for the internal inconsistency errors inconsistent information was inserted in sentences two and four. Thus, children were only required to store and integrate information across one sentence to detect and identify the error. Consequently, it may be the case that the passages used in the current experiment did not have sufficient processing and storage demands to tap children's working memory. Further research in this thesis will explore the relationship between comprehension monitoring and working memory by continuing to include independent measures of working memory, and work presented in Chapters 5 and 6 manipulates the distance between inconsistent information within passages.

3.3. General discussion

In line with previous research, the results of Experiment 1 suggest that there are developmental differences in children's comprehension monitoring ability. Younger children do not engage in accurate monitoring as frequently as older children, and thus are poorer at judging whether or not a text makes sense and correctly identifying errors within a text. In general, children were very good at monitoring their comprehension for nonword errors. However, children's monitoring of general knowledge violation and internal inconsistency errors was less accurate. This may imply that children are proficient at using a lexical standard of evaluation, whereas the use of external consistency and internal consistency standards is still developing in 7- to 10-year-old children. In contrast, it may be that failing to identify an error is a sign of good comprehension. Children may have detected the error and taken remedial action so that it is no longer perceived as problematic. Future experiments in this thesis, will aim to further explore 7- to 10-year-old children's comprehension monitoring of general knowledge violation and internal inconsistency errors. In addition, whilst children were able to detect an error in a text, they did not always go on to accurately identify the error. Future experiments in this thesis will use the real-time methodology of reading times and eye-movements to provide a greater insight into instances when children detect an error but fail to report the error, and whether an error is detected during reading or when prompted by features of the task, such as the instructions delivered by the experimenter.

Although beyond the scope of this thesis, the procedure used in Experiment 1 provides an interesting avenue for future research. Children were instructed to underline any errors in the text and then respond to the question 'did the story make sense?' by ticking a happy or sad face. This task order is in line with a real-time view of comprehension monitoring, in which children monitor their comprehension, evaluate the sense of information and detect errors during reading. However, it may be that children read the entire passage before engaging in monitoring behaviours. Thus, it is possible that the alternate task order could be used in monitoring experiments. Putting children in a positon where they have to decide whether or not a passage makes sense prior to underlining erroneous information may influence monitoring behaviours, for example findings may reveal more accurate identification of error information.

The results of Experiment 2 suggest that reading skills in younger children, specifically reading accuracy, contribute to their comprehension monitoring ability. Thus, it seems that accurate word reading facilitates younger children's comprehension monitoring ability. Perhaps by reducing the demands on their information processing capacity, leaving sufficient resources to undertake monitoring processes such as integration. Surprisingly, independent measures of reading comprehension and working memory did not contribute to comprehension monitoring ability. It can be argued that the comprehension monitoring task was not demanding of working memory processes, so the relationship between comprehension monitoring ability and working memory requires further investigation. Future experiments in this thesis will include an independent measure of working memory and manipulate the working memory demands of the comprehension monitoring task by increasing the distance between inconsistent information within passages.

CHAPTER FOUR

A READING TIME EXPERIMENT INVESTIGATING HOW DIFFERENT INSTRUCTIONS INFLUENCE COMPREHENSION MONITORING

The central aim of this chapter was to investigate whether task instructions influence children's monitoring behaviour. Children completed two conditions with different instructions. In the first, children were not explicitly alerted to the presence of errors. In the second, children were alerted to the presence of errors. Critically, children's reading times were recorded so that their real-time reading behaviour could be assessed.

4.1. Experiment 3

In line with previous research, Experiment 1 demonstrated developmental differences in children's comprehension monitoring. One proposed reason for this difference is an increase in metacognitive skills (e.g., Ehrlich, 1996; see the Literature Review for further information) and this is why instructions were manipulated in the current experiment. Experiment 1 also established that 7- to 8-year-old and 9- to 10-year-old children's monitoring was not at ceiling for either general knowledge violation or internal inconsistency errors. Thus, both error types were included in the current experiment so that children's real-time reading behaviour for these error types could be explored. In addition, to enable comparison with the previous experiments the same age groups (7- to 8-year-olds and 9- to 10-year-olds) participated in the current experiment.

Research has shown that explicit instructions that direct children to search for inserted errors in texts increase the likelihood of error detection and identification (e.g., Baker, 1984a; 1984b). There are several reasons why monitoring performance may increase when children are forewarned that errors are present in texts. One reason is that when children are informed that errors may be present in texts, they set specific goals which foster high levels of evaluation so that their reading involves a more purposeful search for errors (e.g., Baker, 1984a). As a consequence, error detection and identification rates are increased. Also, when children are specifically informed what standard(s) of evaluation to adopt, they are more likely to use the correct standard, and use several different standards in combination to effectively monitor their comprehension (e.g., Baker, 1984b). A further reason is that children may be more likely to report problems when they are informed that texts might contain inserted errors, because they may be more willing to report that the text does not make sense and/or that they are experiencing comprehension problems (e.g., Markman, 1979).

Not all studies have shown that explicit instructions influence the likelihood of error detection and identification (e.g., Garner & Anderson, 1982). The differences between studies may be due to the age of participants and the ease of materials. Studies that have not found significant effects of explicit instruction have included college students and have included materials that may be considered too easy (e.g., Baker & Anderson, 1982). To extend the findings of Experiment 1, the instructions given to children were manipulated in the current experiment: Children were either alerted or not alerted to the presence of inserted errors in texts.

Given the multi-dimensional nature of children's comprehension monitoring, a combination of off-line and real-time measures may be necessary to obtain reliable and valid information about their comprehension monitoring behaviour. As discussed in the Literature Review, it has been suggested that off-line measures of cognitive activities, such as verbal reports and responses to questions, do not necessarily coincide with actual behaviour (e.g., Nisbett & Wilson, 1977). It is possible that children may have monitored their comprehension of the text and detected the error during reading, but failed to notice the internal detection signal or interpret its meaning (e.g., Harris et al., 1981). If so, real-time measures, such as reading times, may reveal longer reading times for problematic text information despite the absence of explicit awareness of the error in terms of correctly answering the question 'did the story make sense'. Thus, real-time measures may provide additional information about children's reading behaviour that is valuable in achieving a true insight into their comprehension monitoring. In addition, the reading time measure is advantageous as it can be used without requesting that readers purposefully search for errors. For example, when readers are not explicitly informed that their task is to detect and identify inserted errors in texts. Therefore, it is less likely to bias reading behaviour. Experiment 3 used a combination of off-line and real-time methods.

Experiment 3 investigated 7- to 8-year-old and 9- to 10-year-old children's ability to monitor their comprehension for general knowledge violation and internal inconsistency errors, when instructions differed in the extent to which they alerted children to the presence of errors in passages. This was achieved by having children complete two conditions with different instructions. In the first, children were not explicitly alerted to the presence of errors: They were told to read carefully in order to answer a comprehension question about each text. In the second, children were alerted to the presence of errors: They were told that the texts may contain errors and their task was to decide whether or not the texts made sense (hereafter referred to as the comprehension question and the sense question instruction conditions, respectively). Both off-line and real-time measures of comprehension monitoring were used: Response accuracy to a comprehension or sense question and reading times determined by the self-paced (sentence-by-sentence) reading task, both described in detail below.

In line with previous research (e.g., Baker, 1984a) and the findings of Experiment 1, it was predicted that in comparison to older children, younger children would make fewer correct responses to a sense question. It was also predicted that children would make more correct responses to a sense question for general knowledge violation passages and fewer correct responses to a sense question for internal consistency passages. In addition, it was predicted that reading times for the target sentence for all passage types (general knowledge violation, internal inconsistency and consistent) would be longer for the sense question instruction than the comprehension question instruction. It is likely that the sense question instruction will encourage children to undertake a more purposeful reading of the text, including additional checks of the sense of sentences, which may result in longer reading times. In line with previous research (e.g., Albrecht & O'Brien, 1993; van der Schoot et al., 2011), it was predicted that reading times would be longer for the target sentence when it contained an error (general knowledge violation or internal inconsistency) than when it was consistent. This pattern would indicate detection of errors in real-time during reading and suggest that comprehension monitoring has taken place. Further, the findings from Experiment 2 were inconclusive about the relationship between comprehension monitoring and word reading, reading comprehension and working memory. So, independent measures of the latter three constructs were included in the current experiment to further explore this relationship.

Method

Participants

Thirty-six children (19 boys, 17 girls) with a mean age of 8;3 years (*SD* = 3 months, range = 7;9 – 8;9 years) and 24 children (14 boys, 10 girls) with a mean age of 10;3 years (*SD* = 2 months, range = 9;9 – 10;8 years) participated. Children had measured word reading (i.e., TOWRE Sight Word and Phonemic Decoding measures) above a standard score of 85. Children who had special educational needs and those who spoke English as a second language were not included in the experiment. Informed consent was obtained from school headteachers and parents, and children assented to participate.

Materials

Passages from Experiment 1 that were considered to be most sensitive to children's comprehension monitoring performance were selected. Selection was based on a passages ranking (sense question response accuracy for each year group) and the experimenters experience of children's interaction with the passages during Experiment 1. For each error type, the middle-ranking passages were selected as representative of passages that did not include an error that was obvious to detect or an error that was particularly difficult to detect. Whereas, the highest-ranking consistent passages were selected as representative of passages that were most consistently identified as not containing information that may be perceived as problematic. There were a total of 28 passages containing errors: 14 for each of the two error types (general knowledge violation and internal inconsistency) and eight consistent passages that did not contain any errors (see Appendix 2.1 for a full list of passages). Fourteen passages remained the same as in Experiment 1 and 22 passages underwent minor revisions (e.g., the name of a character was changed, information was added to a sentence). Revisions were undertaken to address issues children experienced in Experiment 1 (e.g., pronouncing character names, asking questions to clarify information included in a passage). As in Experiment 1, all passages were a similar length (range = 44 – 56, see Table 4.1) and consisted of five sentences.

The mean Flesch-Kincaid Grade readability levels for the general knowledge violation, internal inconsistency and consistent passages were similar to those in Experiment 1 (see Table 4.1). A one-way ANOVA with Flesch-Kincaid Grade Level scores as the dependent variable indicated that there were no significant differences between the Grade Level of the different passage types (general knowledge violation, internal inconsistency and consistent), $F(2,35) < 1$, ns.

Table 4.1. Mean numbers (with standard deviations) of words and Flesch-Kincaid Grade Level for general knowledge violation, internal inconsistency and consistent passages

Passage	**Mean number of Words**	**Mean Flesch-Kincaid Grade Level**
General knowledge violation	48.29 (3.07)	2.77 (.23)
Internal inconsistency	50.00 (3.66)	2.61 (.26)
Consistent	48.25 (3.58)	2.68 (.51)

For the comprehension question instruction, a comprehension question to assess memory for an explicitly stated fact was created for each passage (see Appendix 2.1 for questions). Within the experiment proper, so that a child's ability to store and recall information did not influence response accuracy, the comprehension questions referred to different sentences within the passages. Also, to control for a yes/no bias there were an equal number of questions with yes and no responses.

Self-paced stationary window method

Reading times were collected by means of the self-paced sentence-by-sentence stationary window method (e.g., Just, Carpenter & Woolley, 1982). In this method, participants are presented with passages of text on a computer screen one sentence at a time. When the participant has finished comprehending a sentence, they press a button to move on to the next sentence. In the current experiment, when each new sentence was displayed it replaced the previous sentence, so children could not go back and re-read any previous sentences of a passage. Assuming that information is processed as soon as it is perceived (e.g., Just & Carpenter, 1980), the button pressing latencies (i.e., reading times) in the self-paced reading method reflect the processing of the sentences presented. Sentence reading times were defined as beginning when the sentence was first revealed and lasting until the next button press.

Procedure

Passages were assembled into two lists. Each list consisted of 16 experimental passages: six general knowledge violation, six internal inconsistency and four consistent passages. All passage types were included in each list so that children's sense question accuracy could be compared across the different passage types. Each list also had two practice passages: one internal inconsistency and one general knowledge violation passage. There were different practice passages for each list. Each participant saw both lists: The first list during session one when reading to answer a comprehension question and the second list during session two when reading to answer a sense question.

For each list, the first three passages of the experiment proper were arranged in a constrained order such that a general knowledge violation passage was encountered first, an internal inconsistency passage second and a consistent passage third. The remaining passages were presented in a pseudorandomised order in which no more than two of the same type occurred consecutively within a list. A one-way ANOVA with Flesch-Kincaid Grade Level scores as the dependent variable indicated that there were no significant differences between the Grade Level of passages for the two lists, $F(1,35) < 1$, ns.

Children were tested individually at school in a quiet room, separate from their classroom. They completed the comprehension monitoring tasks and independent measures of word reading (in isolation and in context), reading comprehension and working memory during two 40 minute sessions that took place on different days. In session one, children were instructed that at the end of each story they would be required to answer a question about the story (comprehension question instruction). In session two, children were alerted to the presence of errors. They were instructed that some stories might contain silly mistakes and at the end of each story they would be asked whether the story made sense (sense question instruction).

Each passage was displayed one sentence at a time on a computer screen using E-Prime software (Psychology Software Tools, Pittsburgh, PA). Each sentence was displayed on a new screen. Text was presented in black font on a white background. Children were instructed to read silently at their normal reading pace. At the beginning of each session, children completed the two practice passages to ensure they understood the procedure. In the sense question instruction session, children received feedback on their responses to the sense question to ensure that they were familiar with the two types of error.

Before each passage, children were presented with a screen that prompted them to press the left button of a serial response box to display the passage. When children were ready they pressed the left button of the response box to display the first sentence, and each subsequent sentence, of a passage. Children could not go back and re-read any previous sentences of a passage. After each passage, a comprehension question or the question 'did the story make sense' was displayed on a new screen. Children answered the question by pressing the left button of the response box for 'yes' and the right button of the response box for 'no'. They were not given feedback on the accuracy of their response. Once a response had been given, a new screen was presented that prompted children to press the left button of the response box for the next passage.

For each passage there were two dependent variables: correct sense judgements (accuracy for the 'did the story make sense' question) and sentence reading times. Accuracy for the comprehension questions was not analysed as this was not a dependent variable of interest, rather the comprehension question acted as a means of ensuring that children undertook the reading for meaning task with a purpose that encouraged them to read each passage carefully[4]. Passages and regions of interest (e.g., sentences) were different lengths. To control

---

[4] Note that performance on the comprehension question for the error passages was above 73%, for both age groups, showing that the instructions were effective.

for these differences, prior to analysis reading times per syllable (in milliseconds) were calculated for each sentence. Note that all of the reading time analyses within this thesis transform reading times per sentence (in milliseconds) to reading times per syllable (in milliseconds). The milliseconds per syllable (ms/syllable) measure is an appropriate means of breaking down sentence reading times in experiments in this thesis because syllables are the unit that is most representative of the manner in which 7- to 10-year-old children read (e.g., Adams, 1994). Word and character measures are considered less appropriate as 7- to 10-year-old children do not typically engage in letter-by-letter reading nor do they typically engage in skilled adult sight word vocabulary reading.

Word reading and reading comprehension measures. Children's single word and nonword reading were assessed with the TOWRE, and reading comprehension with the YARC Passage Reading. Full administration details are provided in Chapter 3.

Working memory measures. As in Experiment 2, the Sentence Span and Backward Digit Span tasks from the WMTB-C were administered. However, the number of experimental trials in each task were reduced to decrease the time taken to administer the tasks. Following the practice trials, only the first three experimental trials for each level of task difficulty were administered. The experimenter moved on to the next level of task difficulty when two trials were correctly recalled. Testing was discontinued when children incorrectly recalled two trials within one level of task difficulty.

The trial and span scores were based on a maximum of three rather than six items. Both measures were scored with either 1 for the correct recall of a trial or 0 for failing to correctly recall a trial. Span scores were based on the highest level of task difficulty for which two or more trials had been correctly recalled.

Due to the modifications to the administration procedure of the Sentence Span and Backwards Digits Span tasks, the sample reliability (Cronbach's alpha) for each task was calculated. For this and all other working memory measures in this thesis where the sample reliability is stated, Cronbach's alpha was calculated based on the number of individual trials correct. The Cronbach's alpha was .73 for the Sentence Span task and .79 for the Backward Digit Span task, which was comparable to the reliability reported in the WMTB-C manual for standard administration.

Results

Data filtering

The standard deviation of reading times was examined. No reading times were more than 2.5*SD* from the mean. In line with previous literature (e.g., Albrecht & O'Brien, 1993), reading times less than 500ms were considered as premature or careless responses and excluded from the analysis. If this meant that the reading time for the target sentence (i.e., the sentence containing the error information), or for internal inconsistency passages the first piece of inconsistent information (i.e., sentence two), was omitted because it had a reading time of less than 500ms then the passage was excluded from analysis. This is because it is unlikely that children's reading times and sense question responses will accurately reflect monitoring behaviour if they fail to adequately process critical error information. Additionally, passages with three or more missing sentence reading times were excluded from the analysis**.** Children had to read over half of a passage, three out of five sentences, for it to be included in the analysis. Such a criterion for passage inclusion was used to ensure that children were not repeatedly skipping sentences and had read the passages. This data filtering eliminated approximately 3% of the data analysed. Mean reading times were calculated over the remaining reading times.

Before analysis the Q-Q plots for all measures were examined for each age group separately. The plots indicated that the distribution of the data was acceptable.

Correct sense judgements

As stated in the Method, only correct judgements to the 'did the story make sense' question are reported because correct judgements to the comprehension questions were not a dependent variable of interest. Children did not read the same number of passages for each passage type[5]. Thus, for ease of comparison, the correct sense judgement data was converted into proportions (number of correct sense judgements as a proportion of the total number of passages read for each passage type). The analysis that follows was performed on both untransformed and transformed (arcsine because of proportional data) data sets. The pattern of results was the same for both data sets, thus the untransformed results are reported.

The mean proportion correct for the general knowledge violation, internal inconsistency and consistent passage was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The mean proportions of correct sense judgements made for each passage type by age group are shown in Table 4.2.

Table 4.2. Mean proportions (with standard deviations) of correct judgements by age group and passage type

	Age group		
Passage type	**7- to 8-year-olds (n = 36)**	**9- to 10-year-olds (n = 24)**	**Total**
General knowledge violation	.67 (.21)	.79 (.13)	.73 (.17)
Internal inconsistency	.45 (.29)	.61 (.34)	.53 (.32)
Consistent	.93 (.12)	.98 (.07)	.96 (.10)
Total	.68 (.21)	.79 (.18)	

5 Note this is because there was a different number of error and consistent passages.

There was a significant main effect of age group ($F(1,58) = 8.10$, $p = .01$, $\eta p^2 = .12$) because the older children made more correct sense judgements than the younger children. There was also a main effect of passage type ($F(2,116) = 69.30$, $p < .001$, $\eta p^2 = .54$). Children made more correct sense judgements for the consistent than both the general knowledge violation ($t(59) = 9.69$, $p < .001$, $d = 1.65$) and the internal inconsistency ($t(59) = 10.79$, $p < .001$, $d = 1.81$) passages. In addition, general knowledge violation passages were easier than the internal inconsistency passages ($t(59) = 5.19$, $p < .001$, $d = .78$). The interaction between age group and passage type was not significant ($F(2,116) = 1.25$, $p = .29$, $\eta p^2 = .02$).

Reading times for the target sentence

Analyses comparing the target sentence for the error passages (general knowledge violation and internal inconsistency) with the target sentence for the consistent passages were conducted. The target sentence for the consistent passages was different for each of the error passages. Thus, reading time data is analysed separately for each error type.

Analyses of the general knowledge violation passages. For the general knowledge violation passages, the target sentence was the sentence that contained the knowledge violation. Errors were inserted into either sentence two, three or four. Thus, the mean reading time for these error sentences was used as the reading time for the general knowledge violation target sentence in analyses. The target sentence for the consistent passages was the consistent sentence that related to the inserted knowledge violation errors. Thus, the mean reading time for consistent sentences two, three and four was used as the reading time for the consistent target sentence in analyses.

Two reading time analyses were conducted. The first analysis explores children's reading behaviour as a function of age and instruction, and provides an insight into implicit error detection (i.e., instances when children detected an error, but failed to correctly respond to the sense question). Thus, this analysis compares the comprehension question instruction and sense question instruction conditions and includes all responses (correct and incorrect). The second analysis explores children's reading behaviour as a function of age when the sense question is answered correctly. Thus, this analysis focuses on the sense question instruction condition and includes only correct responses.

Comparison of the two instruction conditions: Analysis of reading times for the target sentence for all responses (correct and incorrect). The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 2 (passage type) x 2 (instruction) ANOVA. The means for the general knowledge violation and consistent passages by age group and instruction are shown in Table 4.3.

There was a significant main effect of age group ($F(1,58) = 22.56$, $p < .001$, $\eta p^2 = .28$) because the older children read the target sentence faster ($M = 307.23$, $SD = 82.46$) than the younger children ($M = 424.86$, $SD = 122.25$). The main effect of instruction was marginally significant ($F(1,58) = 3.56$, $p = .06$, $\eta p^2 = .06$). Children had longer reading times for the target sentence for the sense question instruction ($M = 375.14$, $SD = 99.88$) than the comprehension question instruction ($M = 356.95$, $SD = 104.82$). The main effect of passage type was also marginally significant ($F(1,58) = 3.81$, $p = .06$, $\eta p^2 = .06$). Reading times were longer for the target sentence when it violated general knowledge ($M = 373.34$, $SD = 111.71$) than when it was consistent ($M = 358.75$, $SD = 93.00$). There were no significant interactions (all $Fs < 1$, all $ps > .10$, with the exception of age group · instruction, $F(1,58) = 1.42$, $p = .24$, $\eta p^2 = .02$).

Table 4.3. Mean reading times (ms/syllable) (with standard deviations) for the target sentence by age group, passage type, instruction and question accuracy

		Age group			
		7- to 8-year-olds		9- to 10-year-olds	
		Instruction			
Passage type	**Question accuracy**	**Comprehension**	**Sense**	**Comprehension**	**Sense**
General knowledge violation	**Overall**	418.70 (143.04)	450.21 (127.35)	307.71 (101.00)	316.74 (75.41)
Consistent	**Overall**	401.35 (99.61)	429.19 (118.99)	300.04 (75.63)	304.41 (77.78)
General knowledge violation	**Correct-only**	-	484.51 (155.15)	-	322.48 (78.21)
Consistent	**Correct-only**	-	429.95 (117.67)	-	301.53 (78.45)
Internal inconsistency	**Overall**	432.13 (135.61)	441.76 (128.43)	320.17 (77.27)	320.37 (78.47)
Consistent	**Overall**	387.93 (101.65)	455.07 (135.17)	279.54 (92.22)	311.49 (60.04)
Internal inconsistency	**Correct-only**	-	421.44 (120.14)[a]	-	321.84 (84.26)[b]
Consistent	**Correct-only**	-	453.97 (121.61)[a]	-	315.51 (90.47)[b]

Note. Overall refers to both correct and incorrect responses. Unless indicated in the table by superscript, n = 36 for 7- to 8-year-olds and n = 24 for 9- to 10-year-olds. [a] n = 33, [b] n = 20. For both age groups, missing data in the correct-only analysis is due to children having no correct responses for error passages.

Analysis of the sense question instruction condition: Reading times for the target sentence for correct-only responses. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 2 (passage type) ANOVA. The means for the general knowledge violation and consistent passages by age group are shown in Table 4.3.

A similar pattern of results to the previous analysis for this passage type was found. There was a significant main effect of age group ($F(1,58) = 26.12$, $p < .001$, $\eta p^2 = .31$) because the older children read the target sentence faster ($M = 312.01$, $SD = 78.33$) than the younger children ($M = 457.23$, $SD = 136.41$). There was also a main effect of passage type ($F(1,58) = 9.13$, $p = .01$, $\eta p^2 = .14$). Children had longer reading times for the target sentence when it violated general knowledge ($M = 403.50$, $SD = 116.68$) than when it was consistent ($M = 365.74$, $SD = 98.06$). The interaction between age group and passage type was not significant ($F(1,58) = 1.81$, $p = .18$, $\eta p^2 = .03$).

Analyses of the internal inconsistency passages. For the internal inconsistency passages the target sentence was sentence four. This target sentence was inconsistent with information presented earlier in sentence two of the passages. The target sentence for consistent passages was consistent sentence four, as this related to the sentence in which the internal inconsistency error was inserted. As for the general knowledge violation passages, two analyses are reported: The first compares the two instruction conditions and includes all responses (correct and incorrect); the second focuses on the sense question instruction condition and includes only correct responses.

Comparison of the two instruction conditions: Analysis of reading times for the target sentence for all responses (correct and incorrect). The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 2 (passage type) x 2 (instruction)

ANOVA. The means for the internal inconsistency and consistent passages by age group and instruction are show in Table 4.3.

As for the general knowledge violation passages, there was a significant main effect of age group ($F(1,58) = 21.54$, $p < .001$, $\eta p^2 = .27$) because the older children read the target sentence ($M = 319.56$, $SD = 86.08$) faster than the younger children ($M = 437.15$, $SD = 123.53$). There was also a main effect of instruction ($F(1,58) = 5.51$, $p = .02$, $\eta p^2 = .09$). Children had longer reading times for the target sentence for the sense question instruction ($M = 390.10$, $SD = 104.93$) than the comprehension question instruction ($M = 291.61$, $SD = 104.68$). In addition, there was a main effect of passage type ($F(1,58) = 4.74$, $p = .03$, $\eta p^2 = .08$). Reading times were longer for the target sentence when it was internally inconsistent ($M = 386.91$, $SD = 106.26$) than when it was consistent ($M = 369.80$, $SD = 103.35$).

In contrast to the general knowledge violation passages, there was a significant interaction between passage type and instruction ($F(1,58) = 14.10$, $p < .001$, $\eta p^2 = .20$), which is shown in Figure 4.1. The interaction arose because reading times for the target sentence differed significantly between internal inconsistency and consistent passages for the comprehension question instruction ($t(59) = 4.17$, $p < .001$, $d = .42$), but did not differ significantly for the sense question instruction ($t(59) = 1.09$, $p = .28$, $d = .02$). No other interactions were significant (all $Fs < 1$, $ps > .10$).

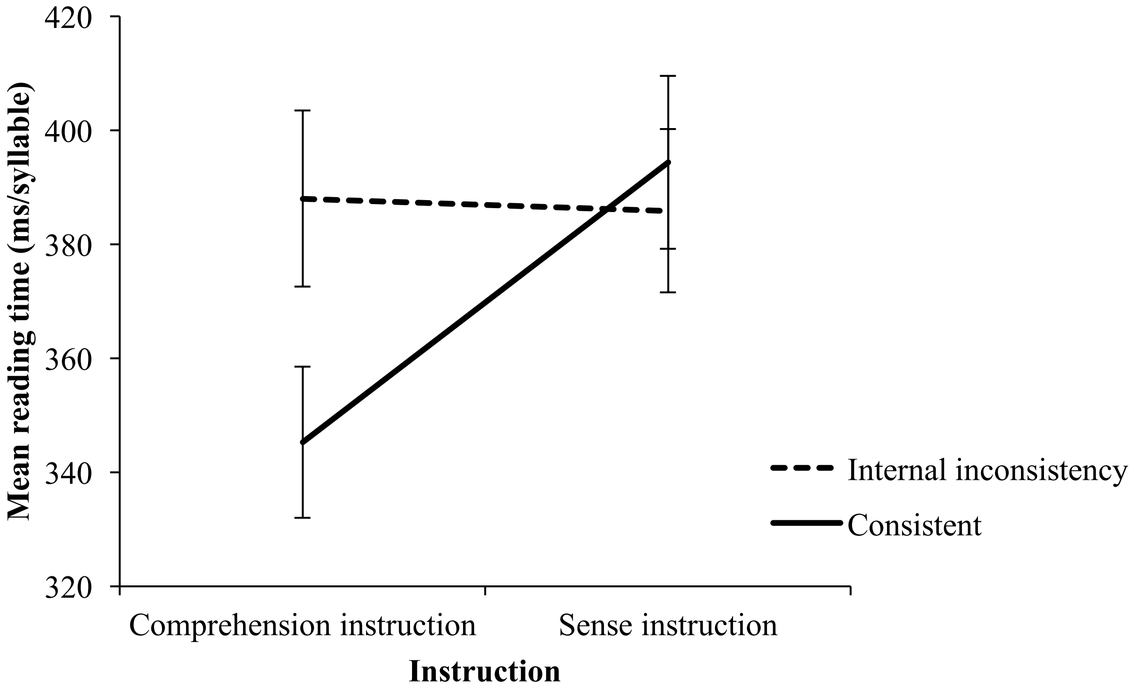

Figure 4.1. Mean reading times (ms/syllable) (+/- standard error) for the target sentence by passage type and instruction

Analysis of the sense question instruction condition: Reading times for the target sentence for correct-only responses. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 2 (passage type) ANOVA. The means for the internal inconsistency and consistent passages by age group are shown in Table 4.3.

Similar to the previous analyses there was a significant main effect of age group ($F(1,51) = 18.19$, $p < .001$, $\eta p^2 = .26$). The older children read the target sentence faster ($M = 318.68$, $SD = 87.37$) than the younger children ($M = 437.71$, $SD = 120.88$). However, in contrast to the previous analyses there was no significant main effect of passage type ($F(1,51) < 1$, ns). Surprisingly, reading times between the inconsistent ($M = 371.64$, $SD = 102.20$) and consistent

(M = 384.74, SD = 106.04) target sentences did not differ significantly. The interaction between age group and passage type was not significant ($F(1,51) = 2.03$, $p = .16$, $\eta p^2 = .04$)[6].

Relationship between comprehension monitoring and word reading, reading comprehension and working memory

Descriptive statistics. Table 4.4 summarises the mean performance for both age groups on independent measures of word reading, reading comprehension and working memory. As mentioned in the method for Experiment 2, it was not appropriate to calculate standard scores for the YARC passage reading comprehension measure because of the procedure followed and the publication of revised manual guidelines. Also, shortened versions of the WMTB-C tasks were administered, so it was not possible to calculate standard scores for the working memory measures. Thus, raw scores are used in analyses. Note that the performance of 7- to 8-year-old and 9- to 10-year-old children on word reading and reading comprehension measures was above a standard score of 85.

6 An analysis comparing reading times for the target sentence for correct vs. incorrect responses to the sense question was conducted. Note that participant numbers were further reduced: 7- to 8-year-olds n = 32 and 9- to 10-year-olds n = 17. Findings were similar to the internal inconsistency vs. consistent analysis. Children had similar reading times for the target sentence when the sense question was answered correctly (M = 377.55, SD = 101.12) and when it was answered incorrectly (M = 383.84, SD = 144.46) ($F(1,47) < 1$, ns).

Table 4.4. Mean raw scores, ability scores and standard scores (with standard deviations and t-tests) for word reading, reading comprehension and working memory measures by age group

Measure	7- to 8-year-olds	9- to 10-year-olds	*t*	*d*
Word reading				
Raw score (maximum = 104)	65.08 (8.22)	74.29 (6.32)	4.64***	1.26
Standardised score	115.08 (8.41)	108.58 (9.47)	2.79**	.73
Nonword reading				
Raw score (maximum = 63)	35.89 (9.70)	45.46 (6.87)	4.18***	1.14
Standardised score	114.97 (11.79)	116.67 (11.04)	< 1	-
Reading rate				
Raw score (seconds)	98.15 (27.55)	105.77 (22.57)	1.13	.30
Ability score	67.33 (8.02)	79.79 (6.07)	6.47***	1.75
Standardised score	109.19 (8.59)	109.56 (8.23)	< 1	-
Word reading accuracy				
Raw score	3.93 (3.07)	3.75 (1.47)	< 1	-
Ability score	52.56 (6.42)	63.77 (3.86)	7.67***	2.12
Standardised score	106.31 (8.33)	108.56 (5.96)	1.15	.31
Reading comprehension				
Raw score (maximum = 8)	3.99 (1.52)	4.58 (1.20)	1.62	.43
Verbal working memory				
Number of items (maximum = 63)	10.78 (5.27)	17.21 (5.22)	4.65***	1.26
Number of trials (maximum = 18)	5.53 (1.96)	7.58 (1.59)	4.28***	1.15
Span (maximum = 6)	1.89 (.78)	2.63 (.65)	3.81***	1.03
Numerical working memory				
Number of items (maximum = 81)	24.69 (10.41)[a]	32.71 (10.74)	2.87**	.76
Number of trials (maximum = 18)	7.06 (2.21)[a]	8.75 (1.96)	3.02**	.81
Span (maximum = 7)	3.46 (.85)[a]	4.00 (.83)	2.42*	.64

Note. Raw scores are presented on the upper row, and where available ability scores on the middle row and standardised scores on the lower row. Unless indicated in the table by superscript, n = 36 for 7- to 8-year-olds and n = 24 for 9- to 10-year-olds. [a] n = 35. Missing data is due one child not completing the Backward Digit Span task. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Interrelations between measures by age group. To explore the relationship between comprehension monitoring ability and word reading, reading comprehension and working memory correlations were calculated for each age group separately. Before analysis a composite measure of word and nonword reading was created from the TOWRE Sight Word reading and TOWRE Phonemic Decoding measures. The composite measure was created based on guidelines that the TOWRE Sight Word reading and TOWRE Phonemic Decoding sub-test scores can be added together to provide an overall composite score (e.g., Boyle & Fisher, 2008; Torgesen et al., 1999) and confirmed by the inspection of the correlations between variables for each age group, which were high (*rs* = .78 and .61; *ps* < .01, for younger and older age groups, respectively). For the verbal and numerical working memory measures, the number of correct items was used in analysis because this can be considered a more sensitive measure of the performance level reached by each child than the memory span (e.g., Conway et al., 2005). The verbal and numerical working memory measures were not combined to form a composite measure so that their relationship with comprehension monitoring could be explored independently. The measure of monitoring performance was accuracy to the sense question for both passage types. Consistent passages were not included in this analysis because performance for 9- to 10-year-old children on the sense question for this passage type was close to ceiling.

For brevity the reader is directed to Tables 4.5 and 4.6, which summarise the correlations for each age group, and only key findings are discussed here. Many correlations did not reach statistical significance (probably because of the small sample size), but we can interpret these in relation to effect sizes where .10 represents a small effect, .30 a moderate effect and .50 a large effect (e.g., Field, 2005). For the 7- to 8-year-old children, there was a negative moderate association between correct responses to the sense question for the general knowledge violation passages and the reading accuracy measure. However, for the younger age group there were no moderate or strong associations between correct responses to the sense question for the internal inconsistency passages and the other measures. For the 9- to 10-year-old children, there was a strong association between correct responses to the sense question for internal inconsistency passages and the reading comprehension measure. However, for the older age group there were no moderate or strong associations between correct responses to the sense question for the general knowledge violation passages and the other measures.

Table 4.5. Correlations between measures: 7- to 8-year-olds

	1	2	3	4	5	6	7
1. Word and nonword reading	-	-.66***	.19	.17	.10[a]	.22	-.02
2. Word reading accuracy	-.66***	-	-.35*	-.18	-.07[a]	-.48**	-.26
3. Reading comprehension	.19	-.35*	-	.32	-.09[a]	.29	.18
4. Verbal working memory	.17	-.18	.32	-	.45**[a]	.25	.12
5. Numerical working memory	.10[a]	-.07[a]	-.09[a]	.45**[a]	-	.05[a]	.19[a]
6. Sense question: General knowledge violation	.22	-.48**	.29	.25	.05 [a]	-	.30
7. Sense question: Internal inconsistency	-.02	-.26	.18	.12	.19 [a]	.30	-

Note. Unless indicated in the table by superscript n = 36. [a] n = 35. Missing data is due to one child not completing the Backward Digit Span task. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 4.6. Correlations between measures: 9- to 10-year-olds

	1	2	3	4	5	6	7
1. Word and nonword reading	-	-.51*	.19	.50*	.22	.01	.22
2. Word reading accuracy	-.51*	-	-.05	-.33	-.28	-.08	-.08
3. Reading comprehension	.19	-.05	-	.26	.16	.14	.58**
4. Verbal working memory	.50*	-.33	.26	-	.61**	.15	.06
5. Numerical working memory	.22	-.28	.16	.61**	-	-.05	.19
6. Sense question: General knowledge violation	.01	-.08	.14	.15	-.05	-	.08
7. Sense question: Internal inconsistency	.22	-.08	.58**	.06	.19	.08	-

Note. Unless indicated in the table by superscript n = 24. [a] n = 35. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

The role of word reading, reading comprehension and working memory in comprehension monitoring ability. To explore the specific role of word reading, reading comprehension and working memory in comprehension monitoring ability, two forward stepwise regressions were conducted with performance on the sense question for general knowledge violation and internal inconsistency passages as the dependent variables, respectively. There were no moderate or strong associations between the word reading, reading comprehension and working memory measures and performance on the sense question for internal inconsistency passages in the 7- to 8-year-old children, and general knowledge passages in the 9- to 10-year-old children. Thus, regressions were only conducted for performance on the sense question for the general knowledge violation passages in the younger age group, and performance on the sense question for the internal inconsistency passages in the older age group. For the reasons mentioned above, consistent passages were not included in analyses. Predictors entered were: word and nonword reading (composite TOWRE Sight Word reading, TOWRE Phonemic Decoding and YARC Passage Reading word reading accuracy measures[7]), reading comprehension, verbal working memory (Sentence Span task item measure) and numerical working memory (Backward Digit Span task item measure).

For general knowledge violation passages, there was no discernible linear relationship between the predictors entered and performance on the sense question. For internal inconsistency passages, the model explained 33% of variance in performance on the sense question ($F(1,23) = 10.95$, $p = .01$). Reading comprehension was the only significant predictor of performance ($\beta = .58$, $p = .01$).

[7] Correlation between measures, *rs* = .61 and .81, *ps* < .01 for younger and older age groups, respectively. See Tables 4.5 and 4.6 for correlations between composite TOWRE word and nonword reading and YARC Passage Reading word reading accuracy measures.

Discussion

The main aim of Experiment 3 was to investigate how different instructions influence 7- to 8-year-old and 9- to 10-year-old children's off-line and real-time comprehension monitoring behaviour for two types of error: general knowledge violations and internal inconsistencies. As predicted, and in line with Experiment 1 (Chapter 3), differences between age groups and error type were found: Older children made significantly more correct sense judgements than younger children, and both age groups made more correct sense judgements for the general knowledge violation passages than the internal inconsistency passages. In an extension to Experiment 1, the instructions presented to children were manipulated and the measurement of reading times was recorded. These yielded novel findings, which are discussed below.

First, the main findings that are convergent with the off-line study (Experiment 1) will be considered. The same pattern of superior error detection for general knowledge violations over internal inconsistencies was apparent and, as before, the older children were more likely to respond that passages containing general knowledge violations did not make sense than passages containing internal inconsistencies. Experiment 3 used a very different paradigm, sentence-by-sentence reading. Thus, it is reassuring that the same basic findings for error type and age are apparent.

Next, the novel findings that came from the use of a real-time task will be discussed. Of interest was the finding that, for both age groups, task instructions influenced real-time reading behaviour. Although reading times for the target sentence were longer for the sense question condition than for the comprehension question condition in general, different patterns of findings in relation to the instructions were apparent for the two error types. For the general knowledge violation passages, the longer reading times for the target sentence in the error passages occurred regardless of task instructions. For the internal inconsistency passages, there was no difference in reading times for consistent and inconsistent target sentences when alerted to the possible presence of errors. In contrast, there was a processing cost for the inconsistent target sentences when children were not alerted to the possible presence of errors, that is in the comprehension question instruction condition.

These findings suggest that instructions that alert children to the likely presence of errors in texts change children's reading behaviour. One likely reason for this is that increased reading time reflects a more purposeful and careful search for errors (e.g., Baker, 1984a). Thus, children may slow down their processing to foster higher levels of evaluation and engage in critical monitoring strategies, such as double-checking the sense of sentences. There are two sources of evidence for this. The first is that there was a general increase in reading times for the target sentences in the sense question instruction condition. The second is that there was no difference between the inconsistent and consistent passages in this condition for the internal inconsistencies. For this error type, the sentence containing the inconsistency was always sentence four, whereas the sentence position varied for the general knowledge violations. Thus, it may be that the absence of a difference between passages that were consistent and those containing an error reflects awareness of this and strategic checking behaviour.

Further, for both the analysis for all responses (and thus regardless of accuracy) and correct-only responses, it was of interest that reading times were generally longer for the target sentence when it contained an error than when it was consistent. This finding suggests that even when children fail to correctly answer the sense question at some level they are detecting the error, and subsequently slowing down their processing rate. If children generate a real-time signal of error detection, why do they indicate that these problematic texts make sense? One possibility is that children's memory for the error information was poor and at the end of the passage they could not remember the information necessary to correctly answer the sense question. However, it is unlikely that this explanation is adequate as passages were short and children did not demonstrate poor performance on the working memory measures. A more probable explanation is that children generated an internal signal of comprehension failure during reading, but failed to notice the signal or interpret its meaning (e.g., Harris et al., 1981).

In sum, the reading time findings provide clear evidence that this was a sensitive measure of children's comprehension monitoring: Older children read the target sentence faster than younger children. This finding was predicted given older children's increased reading fluency and higher information processing capabilities. Also, of interest, there were no interactions between age and either error type or the instruction manipulation, indicating that both age groups approached the task similarly, albeit with different levels of success. As noted in Experiment 1 (Chapter 3), this experiment included more error than consistent passages. To account for this, the sense judgement analysis used proportions and no ceiling effects were no found. Nonetheless, this is a potential limitation which may have biased reading or comprehension monitoring strategy.

Finally, the findings that came from exploring the relationship between comprehension monitoring and independent measures of word reading, reading comprehension and working memory will be considered. For the younger children, word reading accuracy was related to performance on the sense question for general knowledge violation passages. This is perhaps unsurprising as it has been suggested that inaccurate word reading may negatively affect comprehension monitoring by placing too large a demand on children's information processing capacity, leaving little remaining for undertaking important monitoring processes (e.g., Perfetti, 1985). For the older children, reading comprehension was related to, and explained unique variance in, performance on the sense question for internal inconsistency passages. It is likely that this relationship between comprehension monitoring and general reading comprehension reflects the importance of comprehension monitoring skills in constructing an integrated and coherent representation of a texts meaning (e.g., Cain et al., 2004).

CHAPTER FIVE

GENERAL KNOWLEDGE VIOLATIONS: READING TIME EXPERIMENTS INVESTIGATING THE EXPLICITNESS OF ERROR INFORMATION

The main aim of the two experiments reported in this chapter was to investigate whether the explicitness of the general knowledge violation influenced children's error detection and real-time reading behaviour. In addition, the presentation style of text was manipulated (each sentence presented on a separate line, Experiment 4 vs. each sentence presented in continuous text, Experiment 5) to determine if this influenced children's error detection and real-time reading behaviour.

Introduction

Previous experiments in this thesis demonstrated developmental differences in children's correct responses to the sense question: Older children were better at correctly judging whether or not a passage containing a general knowledge violation error made sense than younger children. In contrast, the reading time measure (used in Experiment 3, see Chapter 4 for a full description) suggested that both age groups had similar real-time reading behaviour: Regardless of age, children took longer to read the target sentence when it contained a general knowledge violation than when it did not. Experiment 3 also suggested that real-time reading behaviour was influenced by the instructions given to the children: They spent longer reading the target sentence when provided with instructions that alerted them to the possibility that some texts contained errors. The two experiments reported in this chapter manipulated the explicitness of the general knowledge violation to gain a greater insight into whether this aspect of the error influences the likelihood of detection. Children were provided with instructions that alerted them to the presence of errors in texts. In contrast to Experiment 3, the manner in which text was presented provided children with the opportunity to look back at previous text. To determine whether text presentation influences real-time reading behaviour and the detection of errors, the experiments reported in this chapter contrast two different types of presentation. Successive sentences of a passage either appeared below each other on separate lines (Experiment 4) or from left to right in connected text (Experiment 5). As in Experiment 3, children's reading times were recorded so that their real-time reading behaviour could be assessed. To enable comparison with the previous experiments in this thesis, different samples of the same age groups (7- to 8-year-olds and 9- to 10-year-olds) participated in each experiment.

Research has shown that children's monitoring of information can be influenced by the explicitness of the information that forms the error (e.g., Joseph et al., 2008; Markman, 1979; Rayner, Warren, Juhasz & Liversedge, 2004; Tumner, Nesdale & Pratt, 1983). Here, explicitness is defined as how closely error information is related to the standard state of affairs in terms of world knowledge. For example, wood is typically chopped with an axe. What is considered a related error may be created by substituting 'axe' for 'knife', which can be used for chopping or cutting, but not typically (or successfully) for wood. In contrast, what is considered an unrelated error may be created by substituting 'axe' for 'fork', which cannot be used for chopping. To extend our understanding of the factors that influence children's comprehension monitoring performance, the explicitness of the error information in relation to world knowledge was manipulated in the current experiments.

In addition, to gain a greater insight into monitoring behaviour when children encounter general knowledge violation errors, the presentation style of text was also manipulated in the current experiments. In Experiment 3, the self-paced stationary window reading method was used (see Chapter 4 for a full description). Whilst this method has the advantage of ensuring that reading times are for the sentence presented, it eliminates the possibility of regressions and re-reading because children are unable to look back at previous sentences. It has been suggested that readers tend to look back to earlier text when they encounter syntactic or semantic inconsistencies (e.g., Poynor & Morris, 2003), and in such cases it is possible that an inability to re-read previous text may influence the reader's ability to monitor their comprehension and understand the text (e.g., Just et al., 1982). In order to examine whether providing children with the opportunity to look back and re-read previous text influences their comprehension monitoring, the two experiments reported here presented passages cumulatively. That is, as children progressed through a passage successive sentences were presented and remained presented until the entire passage was displayed.

Further, it should be noted that there are two key methodological differences between the general knowledge violation passages used in the current experiments and those used in previous experiments. The length of passages was increased so that there was always a minimum of two sentences between the sentence containing the general knowledge violation and the final sentence of the passages. This was to enable spillover effects (i.e., the influence of error information from the target sentence on the post-target sentence) to be investigated. Note, spillover effects were not analysed in Experiment 3 because the spillover sentence was also the final sentence of a passage. Thus, spillover effects could not be interpreted in the absence of wrap-up effects.

Previous research has provided support for both the presence and absence of spillover effects in reading time analyses (e.g., Albrecht & O'Brien, 1993). The presence of spillover effects suggests that the target sentence influences the time spent reading the post-target sentence. This may be due to children starting to read the post-target sentence while completing the processing of error information from the target sentence. In contrast, the absence of spillover effects suggests that the target sentence does not influence the reading time of the post-target sentence. This may be because children complete the processing of the error information in the target sentence before moving on to the post-target sentence. To-date this research has focused on internal inconsistency errors. Therefore, it is of interest to investigate whether or not spillover effects are present in general knowledge violation passages.

Second, matched versions of each passage manipulation were created. Whilst there was a high level of control in the passages used in previous experiments (passages had a similar number of words and readability level, and reading time analyses were conducted per syllable), it may be the case that differences between error and consistent passages were due to differences in the passages content (e.g., the ease with which words can be decoded and children's familiarity with the passage topic), rather than the detection and identification of errors. The use of matched passages minimises differences between passages because only a few words are changed across different versions of each passage.

5.1. Experiment 4: Comprehension monitoring of related and unrelated general knowledge violations when successive sentences in a passage are presented on separate lines

It has been suggested that the visual presentation of text across the reading surface can influence a reader's ability to integrate information and establish text coherence (e.g., Just et al., 1982). In Experiment 4, successive sentences in a passage were presented on separate lines of a computer screen. Rather than integrating sentences and considering the general meaning of the passage, it is possible that this separate line presentation will encourage children to adopt a standard of evaluation that focuses their attention on each sentence. This may foster children's monitoring of the sense of each sentence, rather than the passage as a whole, and consequently influence error detection. Also, separate line presentation may make it easier for children to identify the location of problematic information during any re-reading.

In line with previous research (e.g., Baker, 1984a) and the findings of previous experiments, it was predicted that in comparison to older children, younger children would make fewer correct responses to the question 'did the story make sense'. It was also predicted that children would make more correct responses to the sense question for the unrelated general knowledge violation passages than for the related general knowledge violation passages. This finding would reveal information about the nature of the standards of evaluation that children use to monitor their comprehension, and suggest that children use the relatedness of information to world knowledge as a standard for monitoring their comprehension. It may also suggest that children are more likely to adopt standards that have lower processing demands; it may be more cognitively demanding to compare related error information with world knowledge than unrelated error information.

In line with previous research (e.g., Albrecht & O'Brien, 1993; van der Schoot et al., 2011) and the findings of Experiment 3, it was predicted that reading times would be longer for the target sentence when it contained an error (i.e., a general knowledge violation) than when it was consistent. As stated in Chapter 4, this pattern would indicate detection of errors in real-time during reading and suggest that comprehension monitoring has taken place. In addition, it was predicted that reading times would be longer for the target sentence when it contained a related than an unrelated general knowledge violation. In comparison to unrelated error information, it is possible that children will experience greater difficulties in judging whether or not related error information is consistent with their world knowledge. Thus, when children encounter related error information they may be required to undertake additional processing such as double-checking the sense of the information presented in relation to their world knowledge and considering situations and/or experiences in which the information presented may be feasible. It is possible that this additional processing will result in longer reading times.

Method

Participants

Twenty-five children (11 boys, 14 girls) with a mean age of 7;9 years (*SD* = 4 months, range = 7;3 – 8;3 years) and 29 children (10 boys, 19 girls) with a mean age of 9;8 years (*SD* = 4 months, range = 9;3 – 10;4 years) participated. As in Experiment 3, children's word reading (measured with the TOWRE Sight Word and Phonemic Decoding) was above a standard score of 85. Children who had special educational needs and those who spoke English as a second language were not included in the experiment. Informed consent was obtained from school headteachers and parents, and children assented to participate.

Materials

Ten general knowledge violation passages from Experiment 3 were modified and 11 new passages were created for the purposes of this experiment, resulting in 21 eight-sentence passages: 18 experimental passages and three practice passages. All passages were a similar length (*M* = 80.07, *SD* = 3.87, range = 75 – 85), and had a similar Flesch-Kincaid Grade readability[8] (*M* = 2.67, *SD* = .28, range = 2.13 – 3.07). There were three versions of each experimental passage: related general knowledge violation, unrelated general knowledge violation and consistent (see Table 5.1 for an example and Appendix 3.1 for a full list of passages). Relative explicitness was determined by the experimenter and checked during the adult pilot (detailed below). Taking the example passage shown in Table 5.1, wood is typically chopped with an axe. A related error was created by substituting ‘axe’ for ‘knife’, which can be used for chopping or cutting, but not typically (or successfully) for wood. In contrast, an unrelated error was created by substituting ‘axe’ for ‘fork’, which cannot be used for chopping. The substituted items in the related and unrelated errors came from the same category, for example ‘knife’ and ‘fork’ are both items of cutlery. The difference between the passages occurred in a single sentence only. Across the sets of passages, the target sentence occurred six times in each of sentences three, four and five.

[8] Note that here and in future Flesch-Kincaid Grade readability calculations, when preparing the passages to be entered into the formula character names were replaced with the word ‘name’. This was done so that character names did not influence passage readability levels.

Table 5.1. Example of a passage

Joe lives with his family.
Every morning, Joe chops wood for the fire.
Joe always uses a knife/a fork/an axe to chop the wood for the fire.
He has to chop at least three logs.
Otherwise, the wood will run out, and the fire will stop burning.
On school days, Joe has to chop the wood quickly.
If he does not chop the wood quickly, he will be late for school.
Sometimes, when Joe is running late, Dad drives him to school in the car.

Note. Order of manipulation for sentence three: related general knowledge violation, unrelated general knowledge violation and consistent.

As for Experiment 1, for the new and modified passages a background knowledge check of the general knowledge for consistent items was conducted with a separate sample of 30 children (15 7- to 8-year-olds; 15 9- to 10-year-olds). Accuracy for all questions across the two age groups was considered acceptable (average 98%, see Appendix 3.2 for a full list of questions). Following the administration of the background knowledge check, a pilot study with 18 adults confirmed that all errors were identifiable. Any unintended errors that were pointed out by the adult readers were modified before the experiment was conducted with children.

Self-paced moving window method

Reading times were collected by means of the self-paced sentence-by-sentence moving window method (e.g., Just et al., 1982). In this method, participants are presented with passages of text on a computer screen in such a way that each sentence is masked by an ‘x’ and sentences are presented one at a time. When the participant has finished comprehending a sentence, they press a button to move on to the next sentence. In the current experiment, once a sentence had been presented to the participant it remained on the screen so that children could go back and re-read previous sentences of a passage. This is in contrast to Experiment 3, where each new sentence replaced the previous sentence. As in Experiment 3, sentence reading times were defined as beginning when the sentence was first revealed and lasting until the next button press.

Procedure

Passages were assembled into three lists. Each list consisted of 18 experimental passages: six related general knowledge violation, six unrelated general knowledge violation and six consistent passages. Each list also had three practice passages: one related general knowledge violation, one unrelated general knowledge violation and one consistent passage. The same practice passages were used for each list. The different versions of each experimental passage were counterbalanced across the lists so that each version of a given passage occurred only once in each list, and so that six passages had the target sentence in sentence three, six in sentence four and six in sentence five. Following the practice, the first three passages of each list were arranged in a constrained order such that a related general knowledge violation passage was encountered first, an unrelated general knowledge violation passage second and a consistent passage third. The remaining passages were presented in a pseudorandomised order in which no more than two of the same type occurred consecutively within a list. Each participant saw one list.

Children were tested individually at school in a quiet room, separate from their classroom. They completed the word reading (in isolation and in context) and reading comprehension measures in one 15 minute session. In two further 30 minute sessions, they completed the comprehension monitoring task and the working memory measures (described later)[9]. Before the start of the comprehension monitoring task, children were informed that the

[9] Note that the children in the current experiment also participated in Experiment 6 (Chapter 6). Thus, general knowledge violation passages were presented in one session and internal inconsistency passages in the other session. The order of presentation was counterbalanced across participants.

task was designed to examine their understanding of short stories, some of which did not make sense. Children were instructed that sentences would be presented one at a time on the computer screen and that they should read silently at their normal reading pace and prepare to answer the question 'did the story make sense?' at the end of each story.

Each passage was displayed in its entirety on one computer screen, one sentence at a time. Each sentence consisted of one line of text and was presented left-aligned on a different line of the screen. Text was presented in black font on a white background. In line with the self-paced moving window method, until presented, each sentence of the text was masked by an 'x'.

At the start of the comprehension monitoring task, the three practice passages were presented to ensure that children understood the procedure. Children also received feedback on their responses to the sense question to ensure they were familiar with the different passage types. Before each passage, children were presented with a screen that prompted them to press the middle button of a response box to display the passage. When children were ready they pressed the middle button of the response box to display the first sentence, and each subsequent sentence, of a passage. Previous sentences of the passage remained revealed to the children as they progressed through the passage. After each passage, the question 'did the story make sense' was displayed on a new screen. Children answered the question by pressing the left button of the response box for 'yes' and the right button of the response box for 'no'. They were not given feedback on the accuracy of their response. Once a response had been given, a new screen was presented that prompted children to press the middle button of the response box for the next passage.

For each passage there were two dependent variables: Correct sense judgements (accuracy for the 'did the story make sense' question) and sentence reading times in milliseconds per syllable.

Word reading, reading comprehension and working memory measures

Word reading (in isolation) measure. Children's single word and nonword reading were assessed with the TOWRE. Full administration details are provided in Chapter 3.

Word reading (in context) and reading comprehension measures. As in previous experiments, the YARC Passage Reading test was administered. In contrast to the procedure used previously, children read two consecutive passages from Form A at a level indicated by their word reading ability (assessed by the Single Word Reading Test (SWRT); Foster, 2007). Each passage was accompanied by a set of eight comprehension questions. The second passage was selected in line with manual guidelines. This administration procedure allowed standardised scores to be calculated. The reliability coefficients for the grade level passages in Form A range from .75 to .95 for the word reading accuracy and the reading rate measures, and .71 to .77 for the reading comprehension measure (consecutive passage administration).

This experiment included different verbal and numerical working memory measures to the previous experiments. This is because the previous experiments did not find a strong relationship between comprehension monitoring and working memory. One reason for these null findings may be that the working memory measures did not adequately tap this skill in the sample of interest. Alternative sentence span and digit span tasks were selected. Selection was based on including tasks that have been found to reveal a relationship between comprehension and/or comprehension monitoring and working memory (e.g., Oakhill et al., 2005; Seigneuric et al., 2000). These new measures are described below.

Verbal working memory measure. A sentence span task was administered based on the work of Towse, Hitch and Hutton (1998). Children were asked to listen to lists of unrelated sentences (3 – 8 words in length). Immediately after hearing each sentence, children had to supply the missing final word of each sentence (shown in parentheses in the following example), e.g., I can see with my (eyes); a house is made of (bricks) (the processing component). When all of the sentences in a trial had been heard, children had to recall the final word that they supplied for each sentence, in the correct serial order (the storage component). If children did not give the words in the correct serial order, they were encouraged to indicate word position. The task began with three practice trials, each at the easiest difficulty level of two sentences. Appropriate feedback was given on the practice trials, but no feedback was given during the test proper. Difficulty increased to a maximum level of five sentences in a trial. There were three test trials for each level of difficulty (see Appendix 3.3 for a full list of materials). When two trials were incorrectly recalled within one level of difficulty, the experimenter completed administering trials for the current difficulty level then discontinued testing. The sample reliability[10] (Cronbach's alpha) was .59, which was comparable to the reliabilities reported for standardised measures of verbal working memory, such as the Sentence Span task from the WMTB-C.

Numerical working memory measure. A digit span task was administered based on the work of Yuill et al. (1989). Children were asked to read groups of three separate digits out loud (the processing component). When all of the groups of digits had been read, children had to recall the final digit of each group, in the correct serial order (the storage component). For example, for the trial 4 – 8 – 4 followed by 7 – 0 – 1, the correct response would be 4 – 1. If children did not give the digits in the correct serial order, they were encouraged to indicate digit position.

The digit sets were printed in black font on white paper and were viewed through a cardboard 'window'. The viewing window was operated by the experimenter: Once the child

---

10 Here, sample refers to children who participated in Experiments 4 – 7.

had read the first group of three digits, the window was moved to obscure the group and reveal the next group, or after the final group onto blank space. There were six practice trials, two at the beginning of each level of difficulty. Appropriate feedback was given on the practice trials, but no feedback was given during the test proper. Difficulty increased to a maximum level of four groups of digits in a trial. There were three test trials for each level of difficulty (see Appendix 3.3 for a full list of materials). When two trials were incorrectly recalled within one level of difficulty, the experimenter completed administering trials for the current difficulty level then discontinued testing. The sample reliability (Cronbach's alpha) was .70, which was in line with previously reported reliabilities for this task (e.g., Oakhill et al., 2005; Seigneuric et al., 2000).

The verbal and numerical working memory measures were scored in the same manner. For the item and trial measures, children were scored with either 1 for correct recall or 0 for incorrect recall. Span scores were based on the highest level of task difficulty for which two or more trials had been correctly recalled. For each measure, there were three dependent variables: The number of items that were correctly recalled, the number of trials that were correctly recalled and the maximum level of difficult (i.e., span) reached.

Results

Data filtering

Reading times per syllable (in milliseconds) were calculated for each sentence. The standard deviation of reading times was examined. No reading times were more than 2.5*SD* from the mean. As in Experiment 3 (Chapter 4), reading times that were less than 500ms were excluded from analysis. In addition, the reading times for sentences that followed those that were less than 500ms were also excluded from analysis. This is because sentences remained revealed to the reader, so it is likely that the reading time for a sentence that follows a sentence with a reading time of less than 500ms will reflect the processing of both sentences[11]. In addition, passages with five or more missing sentence reading times were excluded from the analysis. This data filtering eliminated less than 2.5% of the data analysed. Mean reading times were calculated over the remaining reading times. Before analysis the Q-Q plots for all measures were examined for each age group separately. The plots indicated that the distribution of the data was acceptable.

Correct sense judgements

The mean number correct for the related general knowledge violation, unrelated general knowledge violation and consistent passages was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The mean numbers of correct sense judgements made for each passage type by age group are shown in Table 5.2.

Table 5.2. Mean numbers (with standard deviations) of correct judgements by age group and passage type

	Age group		
Passage type	**7- to 8-year-olds (n = 25)**	**9- to 10-year-olds (n = 29)**	**Total**
Related general knowledge violation (maximum = 6)	4.38 (1.20)	4.62 (1.02)	4.50 (1.11)
Unrelated general knowledge violation (maximum = 6)	4.56 (1.14)	5.40 (.78)	4.98 (1.08)
Consistent (maximum = 6)	5.22 (1.26)	5.58 (.60)	5.40 (.93)
Total (maximum = 6)	4.72 (1.20)	5.20 (.80)	

[11] Note that in contrast to the data filtering undertaken in Experiment 3, in this and all subsequent reading time experiments in this thesis if filtering meant that the reading time for the target sentence was omitted the passage was *not* excluded from analysis. This is because sentences remained revealed to the reader, so they were provided with the opportunity to look back and read this information.

There was a significant main effect of age group ($F(1,52) = 6.79$, $p = .01$, $\eta p^2 = .12$) because the older children made more correct sense judgements that the younger children. There was also a main effect of passage type ($F(2,104) = 12.91$, $p < .001$, $\eta p^2 = .20$). Children made more correct sense judgements for the consistent than both the related general knowledge violation ($t(53) = 4.69$, $p < .001$, $d = .88$) and the unrelated general knowledge violation passages ($t(53) = 2.68$, $p = .01$, $d = .42$). In addition, children made more correct sense judgements for the unrelated general knowledge violation passages than the related general knowledge violation passages ($t(53) = 2.68$, $p = .01$, $d = .44$). The interaction between age group and passage type was not significant ($F(2,104) = 1.46$, $p = .24$, $\eta p^2 = .03$).

Reading times

Two analyses comparing the target sentence for the three passage types were conducted. The first analysis explores children's implicit error detection and includes all responses (correct and incorrect) to the sense question. The second analysis explores children's reading behaviour when the sense question was answered correctly and includes correct-only responses to the sense question.

Analysis of reading times for the target sentence for all responses (correct and incorrect) to the sense question. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the related general knowledge violation, unrelated general knowledge violation and consistent passages by age group are shown in Table 5.3.

Table 5.3. Mean reading times (ms/syllable) (with standard deviations) for the target sentence and post-target sentence by age group, passage type and question accuracy

		Age group			
		7- to 8-year-olds (n = 25)		9- to 10-year-olds (n = 29)	
Passage type	**Question accuracy**	**Target sentence**	**Post-target sentence**	**Target sentence**	**Post-target sentence**
Related general knowledge violation	**Overall**	487.93 (180.51)	484.95 (133.28)	326.21 (100.03)	331.07 (123.17)
	Correct-only	517.07 (200.32)	496.41 (147.75)	337.04 (114.80)	337.51 (136.80)
Unrelated general knowledge violation	**Overall**	490.71 (169.28)	513.93 (157.75)	334.75 (128.71)	341.82 (143.28)
	Correct-only	474.02 (175.55)	507.53 (171.06)	330.64 (138.07)	333.40 (142.87)
Consistent	**Overall**	382.63 (107.28)	503.59 (181.09)	293.86 (96.12)	334.53 (120.54)
	Correct-only	391.92 (165.52)	499.62 (177.77)	293.13 (107.55)	330.77 (123.46)

Note. Overall refers to both correct and incorrect responses.

There was a significant main effect of age group ($F(1,52) = 18.63$, $p < .001$, $\eta p^2 = .28$) because the older children read the target sentence faster ($M = 318.27$, $SD = 108.29$) than the younger children ($M = 453.76$, $SD = 152.36$). There was also a main effect of passage type ($F(2,104) = 14.03$, $p < .001$, $\eta p^2 = .21$). Children had longer reading times for the target sentence when it contained a related general knowledge violation ($M = 407.07$, $SD = 140.27$) and an unrelated general knowledge violation ($M = 412.73$, $SD = 149.00$) than when it was consistent ($M = 338.25$, $SD = 101.70$). In addition, children had similar reading times for the target sentence when it contained a related and an unrelated general knowledge violation.

These two main effects were qualified by a significant interaction ($F(2,104) = 3.36$, $p = .04$, $\eta p^2 = .06$), which is shown in Figure 5.1. The interaction arose because of a magnitude difference in the mean reading times for the three passages types within each age group. For both age groups, reading times for the target sentence differed significantly between related general knowledge violation and consistent passages ($t(24) = 3.30$, $p = .01$, $d = .71$; $t(28) = 2.69$, $p = .01$, $d = .33$, for younger and older age groups, respectively) and unrelated general knowledge violation and consistent passages ($t(24) = 4.14$, $p < .001$, $d = .76$; $t(28) = 3.03$, $p = .01$, $d = .36$). However, reading times for the target sentence did not differ significantly between related and unrelated general knowledge violation passages (both *ts* < 1, ns).

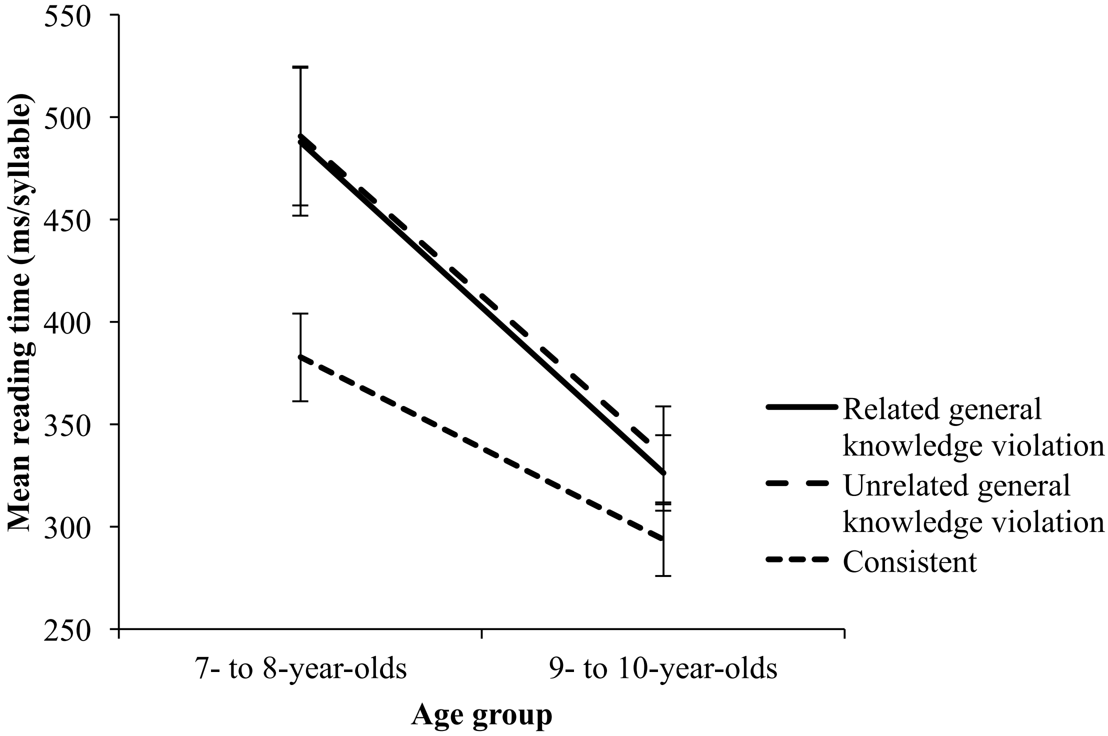

Figure 5.1. Mean reading times (ms/syllable) (+/- standard error) for the target sentence by age group and passage type

Analysis of reading times for the target sentence for correct-only responses to the sense question. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the related general knowledge violation, unrelated general knowledge violation and consistent passages by age group are shown in Table 5.3.

Similar to the previous analysis there was a significant main effect of age group ($F(1,52) = 16.75$, $p < .001$, $\eta p^2 = .24$) because the older children read the target sentence faster ($M = 320.27$, $SD = 120.14$) than the younger children ($M = 461.00$, $SD = 180.46$). There was also a main effect of passage type ($F(2,104) = 9.50$, $p < .001$, $\eta p^2 = .15$). Children had longer reading times for the target sentence when it contained a related general knowledge violation ($M = 427.06$, $SD = 157.56$) and an unrelated general knowledge violation ($M = 402.33$, $SD = 156.81$) than when it was consistent ($M = 342.53$, $SD = 136.54$) ($t(53) = 3.92$, $p < .001$, $d = .57$ for related general knowledge violation and consistent passages; $t(53) = 3.20$, $p = .01$, $d = .41$ for unrelated general knowledge violation and consistent passages). Children had similar reading times for the target sentence when it contained a related or unrelated general knowledge violation ($t(53) = 1.11$, $p = .27$, $d = .16$). In contrast to the previous analysis, the interaction between age group and passage type was not significant ($F(2,104) = 2.08$, $p = .13$, $\eta p^2 = .04$).

Analysis of spillover effects. Reading times for the sentence that followed the target sentence (post-target sentence) for all responses (correct and incorrect) and correct-only responses to the sense question were analysed to explore possible spillover effects. As the general knowledge violation errors were inserted into either sentence three, four or five, the post-target sentence was sentence four, five or six depending on the position of the inserted error.

The mean reading time for the post-target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the related general knowledge violation, unrelated general knowledge violation and consistent passages by age group are shown in Table 5.3.

There was a significant main effect of age group ($F(1,52) = 22.11$, $p < .001$, $\eta p^2 = .30$; $F(1,52) = 21.57$, $p < .001$, $\eta p^2 = .29$, for all and correct-only responses analyses, respectively) because the older children read the post-target sentence faster than the younger children. There was no significant main effect of passage type in either analysis (both $Fs < 1$, ns). Further, the interaction between age group and passage type was not significant (both $Fs < 1$, ns).

Summary

The main aim of Experiment 4 was to investigate how the explicitness of error information influenced 7- to 8-year-old and 9- to 10-year-old children's comprehension monitoring of general knowledge violation errors when successive sentences in a passage were presented on separate lines. As predicted, and in line with previous experiments in this thesis, differences between age groups were found: Older children made significantly more correct sense judgements than younger children. In an extension of previous experiments, the explicitness of error information was manipulated. These novel findings are outlined below.

Children made more correct sense judgements for the consistent than both the related and unrelated general knowledge violation error passages. Also, in line with predictions children made more correct sense judgements for the unrelated general knowledge violation passages than the related general knowledge violation passages. The reading time findings revealed that children had longer reading times for the target sentence when it contained a general knowledge violation than when it was consistent. This pattern was found in both the analysis for all responses (and thus regardless of accuracy) and also for correct-only responses to the sense question. Thus, suggesting that children detected the general knowledge violation errors. However, in contrast to predictions children's reading times for the target sentence were similar when it contained either a related or an unrelated general knowledge violation. In addition, children's reading times for the post-target sentence were similar for all passage types (related general knowledge violation, unrelated general knowledge violation and consistent).

It is possible that these findings are due to the separate line presentation style of the passages. Separate line presentation may have encouraged children to focus their attention on each sentence. Consequently, children may have been more likely to adopt a standard of evaluation that focuses on monitoring the sense of each sentence in isolation, rather than integrating sentences and considering the general meaning of the passage. Also, separate line presentation may have made it easier for children to locate problematic information during any re-reading. These factors may have fostered children's monitoring performance, perhaps resulting in more correct error detections. To explore this, Experiment 5 was conducted to investigate how the explicitness of error information influenced children's comprehension monitoring of general knowledge violation errors when successive sentences in a passage were presented in connected text.

5.2. Experiment 5: Comprehension monitoring of related and unrelated general knowledge violations when successive sentences in a passage are presented in connected text

In contrast to Experiment 4, successive sentences in passages in the current experiment were presented in connected text, from left to right on each line of a computer screen. Because sentences are linked together in connected text, it is possible that this presentation style may encourage children to adopt a standard of evaluation that is different to that encouraged by separate line presentation. Children may be more likely to integrate information across sentences and focus on the general meaning of a passage, rather than monitoring the sense of individual sentences. These standards may influence monitoring behaviour in different ways. Experiment 5 was conducted to explore the influence of connected text presentation on children's comprehension monitoring and further explore the findings from the previous experiment. As before, main effects of age group and passage type were predicted.

Method

In Experiment 5, the word reading, reading comprehension and working memory measures and procedure were the same as the previous experiment. The only differences between the two experiments were the participants, minor revisions to some of the comprehension monitoring materials and the presentation of materials. Here, only these differences are described.

Participants

Twenty-two children (10 boys, 12 girls) with a mean age of 8;2 years (*SD* = 4 months, range = 7;9 – 8;8 years) and 23 children (12 boys, 11 girls) with a mean age of 10;3 years (*SD* = 3 months, range = 9;9 – 10;9 years) participated. As in previous experiments, children's word reading (measured with the TOWRE Sight Word and Phonemic Decoding) was above a standard score of 85. Children who had special educational needs and those who spoke English as a second language were not included in the experiment. Informed consent was obtained from school headteachers and parents, and children assented to participate.

Materials

Of the 21 passages, 16 remained the same as the previous experiment. Five passages were revised to accommodate the new presentation style (see Appendix 3.4 for a list of passages). All passages were a similar length (*M* = 79.70, *SD* = 3.91, range = 75 – 85) and had a similar Flesch-Kincaid Grade readability (*M* = 2.64, *SD* = .32, 1.98 – 3.07).

Procedure

All passages were presented in a pseudorandomised order in which no more than two of the same type occurred consecutively within a list. Each passage was displayed in its entirety on one computer screen, one sentence at a time. Each sentence was presented in its naturally occurring position from left to right on each line of the screen, with successive lines appearing below each other in connected text. In line with the self-paced moving window method, until presented each character (with the exception of spaces and punctuation) of the text was masked by an 'x'.

Results

Reading times per syllable (in milliseconds) were calculated for each sentence. Data filtering and preparation was the same as in the previous experiment. This data filtering eliminated approximately 3% of the data analysed. No reading times were more than 2.5*SD* from the mean. Before analysis the Q-Q plots for all measures were examined for each age group separately. The plots indicated that the distribution of the data was acceptable.

Correct sense judgements

The mean number correct for the related general knowledge violation, unrelated general knowledge violation and consistent passages was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The mean numbers of correct sense judgements made for each passage type by age group are shown in Table 5.4.

Table 5.4. Mean numbers (with standard deviations) of correct judgements by age group and passage type

	Age group		
Passage type	**7- to 8-year-olds (n = 22)**	**9- to 10-year-olds (n = 23)**	**Total**
Related general knowledge violation (maximum = 6)	3.86 (1.55)	4.82 (1.41)	4.34 (1.48)
Unrelated general knowledge violation (maximum = 6)	4.86 (1.17)	5.30 (.70)	5.08 (.94)
Consistent (maximum = 6)	5.27 (.99)	5.30 (.87)	5.29 (.93)
Total (maximum = 6)	4.66 (1.24)	5.14 (.99)	

The pattern of results was similar to that found for the previous experiment. As before, there was a significant main effect of age group ($F(1,43) = 4.37$, $p = .04$, $\eta p^2 = .09$) because the older children made more correct sense judgements than the younger children. As before, there was also a main effect of passage type ($F(2,86) = 9.99$, $p < .001$, $\eta p^2 = .19$). However, the pattern of difference was not the same as that found in Experiment 4. As in Experiment 4, children made more correct sense judgements for the consistent than the related general knowledge violation passages ($t(44) = 3.57$, $p = .01$, $d = .77$), and more correct sense judgements for the unrelated than the related general knowledge violation passages ($t(44) = 3.44$, $p = .01$, $d = .60$). In contrast to Experiment 4, children made a similar number of correct sense judgements for the consistent and the unrelated general knowledge violation passages ($t(44) = 1.03$, $p = .31$, $d = .22$). As before, the interaction between age group and passage type was not significant ($F(2,86) = 2.20$, $p = .11$, $\eta p^2 = .05$).

Reading times

As in the previous experiment, two analyses comparing the target sentence for the related general knowledge violation, unrelated general knowledge violation and consistent passages are reported: The first includes all responses (correct and incorrect) to the sense question; the second includes correct-only responses to the sense question.

Analysis of reading times for the target sentence for all responses (correct and incorrect) to the sense question. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the related general knowledge violation, unrelated general knowledge violation and consistent passages by age group are shown in Table 5.5.

The pattern of results for the main effects of age group and passage type were the same as the previous experiment. There was a significant main effect of age group ($F(1,43) = 5.50$, $p = .02$, $\eta p^2 = .11$) because the older children read the target sentence faster ($M = 380.12$, $SD = 139.48$) than the younger children ($M = 473.52$, $SD = 152.47$). There was also a main effect of passage type ($F(2,86) = 6.53$, $p = .01$, $\eta p^2 = .13$). Children had longer reading times for the target sentence when it contained a related general knowledge violation ($M = 445.51$, $SD = 147.84$) and an unrelated general knowledge violation ($M = 442.34$, $SD = 160.00$) than when it was consistent ($M = 392.62$, $SD = 130.10$) ($t(44) = 3.03$, $p = .01$, $d = .38$ for related general knowledge violation and consistent passages; $t(44) = 3.34$, $p = .01$, $d = .34$ for unrelated general knowledge violation and consistent passages). In addition, children had similar reading times for the target sentence when it contained a related and an unrelated general knowledge violation ($t(44) < 1$, ns). In contrast to Experiment 4, the interaction between age group and passage type was not significant ($F(2,86) < 1$, ns).

Table 5.5. Mean reading times (ms/syllable) (with standard deviations) for the target sentence and post-target sentence by age group, passage type and question accuracy

		Age group			
		7- to 8-year-olds (n = 22)		9- to 10-year-olds (n = 23)	
Passage type	**Question accuracy**	**Target sentence**	**Post-target sentence**	**Target sentence**	**Post-target sentence**
Related general knowledge violation	**Overall**	496.64 (176.44)	509.09 (181.33)	394.37 (119.23)	382.20 (121.11)
	Correct-only	502.74 (169.71)	522.97 (213.03)	396.47 (140.27)	366.07 (126.21)
Unrelated general knowledge violation	**Overall**	490.33 (171.81)	504.24 (162.61)	394.35 (148.19)	380.49 (124.63)
	Correct-only	494.27 (181.48)	493.32 (148.33)	392.15 (146.78)	380.31 (123.92)
Consistent	**Overall**	433.60 (109.17)	492.38 (144.34)	351.63 (151.02)	399.83 (134.46)
	Correct-only	429.11 (108.67)	500.16 (142.75)	347.93 (144.78)	392.31 (133.08)

Note. Overall refers to both correct and incorrect responses.

Analysis of reading times for the target sentence for correct-only responses to the sense question. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the related general knowledge violation, unrelated general knowledge violation and consistent passages by age group are shown in Table 5.5.

The pattern of results was the same as that found for the previous experiment and the analysis above for all (correct and incorrect) responses to the sense question. There was a significant main effect of age group ($F(1,43) = 5.66$, $p = .02$, $\eta p^2 = .12$) because the older children read the target sentence faster ($M = 378.85$, $SD = 143.94$) than the younger children ($M = 475.37$, $SD = 153.29$). There was also a main effect of passage type ($F(2,86) = 8.32$, $p < .001$, $\eta p^2 = .16$). Children had longer reading times for the target sentence when it contained a related general knowledge violation ($M = 449.61$, $SD = 154.99$) and an unrelated general knowledge violation ($M = 443.21$, $SD = 164.13$) than when it was consistent ($M = 388.52$, $SD = 126.73$) ($t(44) = 3.40$, $p = .01$, $d = .43$ for related general knowledge violation and consistent passages; $t(44) = 3.69$, $p = .01$, $d = .37$ for unrelated general knowledge violation and consistent passages). Children had similar reading times for the target sentence when it contained a related and an unrelated general knowledge violation ($t(44) < 1$, ns). The interaction between age group and passage type was not significant ($F(2,86) < 1$, ns).

Analysis of spillover effects. Separate analyses for the post-target sentence for all responses (correct and incorrect) and correct-only responses to the sense question were conducted. The mean reading time for the post-target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the related general knowledge violation, unrelated general knowledge violation and consistent passages by age group are shown in Table 5.5.

The pattern of results was the same as that found for the previous experiment. There was a significant main effect of age group ($F(1,43) = 8.08$, $p = .01$, $\eta p^2 = .16$; $F(1,43) = 9.30$, $p = .01$, $\eta p^2 = .18$, for all and correct-only responses analyses, respectively) because the older children read the post-target sentence faster than the younger children. There was no significant main effect of passage type (both *Fs* < 1, ns). Further, the interaction between age group and passage type was not significant ($F(2,86) < 1$, ns; $F(2,86) = 1.57$, $p = .22$, $\eta p^2 = .04$).

Summary

The main aim of Experiment 5 was to investigate how the explicitness of error information influenced 7- to 8-year-old and 9- to 10-year-old children's comprehension monitoring of general knowledge violation errors when successive sentences in a passage were presented in connected text. As predicted, and in line with the findings of Experiment 4, differences in age groups were found: Older children made significantly more correct sense judgements than younger children. Novel findings are outlined below.

In general, children made more correct sense judgements for consistent than general knowledge violation error passages. However, in contrast to the findings for Experiment 4, children made a similar number of correct sense judgements for consistent and unrelated general knowledge violation passages. Similar to the findings for Experiment 4, children made more correct sense judgements for the unrelated general knowledge violation passages than the related general knowledge violation passages.

The reading time findings were the same as those for Experiment 4. Children had longer reading times for the target sentence when it contained a general knowledge violation than when it was consistent. This pattern was found in both the analysis for all responses (and thus regardless of accuracy) and also for correct-only responses to the sense question. As before, this finding suggests that children detected the general knowledge violation errors. In contrast to predictions, children's reading times for the target sentence were similar when it contained a related or an unrelated general knowledge violation. In addition, children's reading times for the post-target sentence were similar for all passage types (related general knowledge violation, unrelated general knowledge violation and consistent).

5.3. The relationship between comprehension monitoring and word reading, reading comprehension and working memory

The relationship between comprehension monitoring and word reading, reading comprehension and working memory were explored. Because children in Experiments 4 and 5 completed the same measures, and sense judgement analyses revealed essentially the same pattern of main effects, for reasons of power data from the two experiments were combined for the analyses that follow.

Descriptive statistics

Table 5.6 summarises the performance of the two age groups on the independent measures of word reading, reading comprehension and working memory for Experiments 4 and 5. Because the working memory tasks were not standardised measures, raw and ability scores[12] are used in analyses. Note that the performance for both age groups on the word reading and reading comprehension measures was equal to or above a standard score of 85.

[12] Ability scores are used in analysis for the YARC Passage Reading measures. This is because children read different passages, so their raw scores cannot be compared directly. Raw scores were converted to ability scores, which are estimates of a child's level on the reading skill measured. The ability score reflects both the raw score and the difficulty of the items that have been administered.

Table 5.6. Mean raw scores, ability scores and standard scores (with standard deviations and t-tests) for word reading, reading comprehension and working memory measures by age group

Measure	**7- to 8-year-olds (n = 47)**	**9- to 10-year-olds (n = 52)**	***t*(98)**	***d***
Word reading				
Raw score (maximum = 104)	63.09 (7.80)	70.88 (7.45)	5.09***	1.02
Standardised score	115.51 (8.77)	107.98 (10.25)	3.91***	.79
Nonword reading				
Raw score (maximum = 63)	34.98 (9.65)	41.62 (7.56)	3.83***	.77
Standardised score	116.04 (10.90)	112.90 (10.78)	1.44	.29
Reading rate				
Raw score (seconds)	101.10 (28.73)	108.81 (22.87)	1.48	.30
Ability score	66.96 (9.09)	78.29 (8.06)	6.57***	1.32
Standardised score	111.89 (9.47)	109.35 (11.33)	< 1	-
Word reading accuracy				
Raw score	4.30 (2.41)	3.36 (2.09)	2.08*	.42
Ability score	54.00 (7.90)	63.67 (7.58)	6.57***	1.25
Standardised score	110.19 (10.04)	110.19 (10.89)	< 1	-
Reading comprehension				
Raw score (maximum = 8)	4.79 (1.16)	5.13 (1.11)	1.52	.30
Ability score	55.74 (7.25)	63.06 (7.12)	5.06***	1.02
Standardised score	106.53 (8.86)	105.92 (10.55)	< 1	-
Verbal working memory				
Number of items (maximum = 42)	12.40 (5.16)	18.19 (6.57)	4.84***	.98
Number of trials (maximum = 12)	3.70 (1.43)	5.15 (1.50)	4.92***	.99
Span (maximum = 5)	2.19 (.58)	2.67 (.62)	4.00***	.80
Numerical working memory				
Number of items (maximum = 27)	12.87 (5.39)	18.75 (5.34)	5.45***	1.10
Number of trials (maximum = 9)	4.06 (1.75)	5.67 (1.64)	4.72***	.95
Span (maximum = 4)	2.26 (.64)	2.94 (.67)	5.20***	1.04

Note. Raw scores are presented on the upper row, and where available ability scores on the middle row and standardised scores on the lower row. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Interrelations between measures

To explore the relationship between comprehension monitoring ability and the word reading, reading comprehension and working memory measures correlations were calculated for each age group separately. As in Experiment 3 (Chapter 4), before analysis a composite measure of word and nonword reading was created from the TOWRE Sight Word reading and TOWRE Phonemic Decoding measures (*rs* = .82 and .69; *ps* < .001, for younger and older age groups, respectively). Ability scores for the YARC Passage Reading accuracy and YARC Passage Reading comprehension measures and the number of correct items for the memory measures were used in analysis. Whilst there were moderate associations between the verbal and numerical working memory measures in both age groups, a composite measure was not created so that the independent relations for each measure could be explored. Measures of monitoring performance were accuracy to the sense question for the related and unrelated general knowledge violation passages. Consistent passages were not included in analysis because performance was close to ceiling for the oldest age group.

For brevity the reader is directed to Tables 5.7 and 5.8, which summarise the correlations for each age group, and only key findings are discussed here. Many correlations did not reach statistical significance (probably because of the small sample size). However, the correlations can be interpreted in relation to effect sizes. For the 7- to 8-year-old children, there was a small association between correct responses to the sense question for related general knowledge violations and the word reading accuracy measure. For the younger age group, there were also moderate associations between correct responses to the sense question for unrelated general knowledge violations and the word and nonword reading, word reading accuracy and numerical working memory measures. For the 9- to 10-year-old children, there were no moderate or strong associations between correct responses to the sense question for either of the general knowledge violation passages and the other measures.

Table 5.7. Correlations between measures: 7- to 8-year-olds (n = 47)

	1	2	3	4	5	6	7
1. Word and nonword reading	-	.69***	.08	.29*	.22	.23	.40**
2. Word reading accuracy	.69***	-	.42**	.30*	.16	.29$^{(*)}$	.43**
3. Reading comprehension	.08	.42**	-	.16	.13	.03	.23
4. Verbal working memory	.29*	.30*	.16	-	.39**	.08	.21
5. Numerical working memory	.22	.16	.13	.39**	-	.13	.33*
6. Sense question: Related general knowledge violation	.23	.29$^{(*)}$	.03	.08	.13	-	.48**
7. Sense question: Unrelated general knowledge violation	.40**	.43**	.23	.21	.33*	.48**	-

Note. $^{(*)}p = .05$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 5.8. Correlations between measures: 9- to 10-year-olds (n = 52)

	1	2	3	4	5	6	7
1. Word and nonword reading	-	.59***	.38**	.31	.45**	.05	.10
2. Word reading accuracy	.59***	-	.46**	.37**	.36**	.02	.01
3. Reading comprehension	.38**	.46**	-	.26	.25	.21	.03
4. Verbal working memory	.31*	.37**	.26	-	.45**	.19	.14
5. Numerical working memory	.45**	.36**	.25	.45**	-	.09	.15
6. Sense question: Related general knowledge violation	.05	.02	.21	.19	.09	-	-.02
7. Sense question: Unrelated general knowledge violation	.10	.01	.03	.14	.15	-.02	-

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

The role of word reading, reading comprehension and working memory in comprehension monitoring ability

To explore the specific role of word reading, reading comprehension and working memory in comprehension monitoring ability, two forward stepwise regressions were conducted with performance on the sense question for related and unrelated general knowledge violation passages as the dependent variables, respectively. Because there were no moderate or strong associations between the monitoring measures and the other measures in the older age group, regressions were only conducted for the younger age group. For the reasons mentioned above, consistent passage performance was not included in these analyses. Predictors entered were: word and nonword reading (composite TOWRE Sight Word reading, TOWRE Phonemic Decoding and YARC Passage Reading word reading accuracy measures [13]), reading comprehension, verbal working memory (Sentence Span task item measure) and numerical working memory (Digit Span task item measure).

For related general knowledge violations, there was no discernible linear relationship between the predictors entered and performance on the sense question. For unrelated general knowledge violations, the model explained 18% of variance in performance on the sense question ($F(1,45) = 9.58$, $p = .01$). Word and nonword reading was the only significant predictor of performance ($\beta = .42$, $p = .01$).

5.4. Discussion

First, findings for the sense question will be considered. A similar pattern of off-line findings was found in both experiments. This suggests that presentation style did not influence children's ability to correctly judge whether or not a passage made sense. Both experiments

[13] Correlation between measures *rs* = .82 and .71, *ps* < .001 for younger and older age groups, respectively. See Tables 5.7 and 5.8 for correlations between composite TOWRE word and nonword reading and YARC Passage Reading word reading accuracy measures.

revealed developmental differences: Older children made more correct sense judgements than younger children. This is in line with the literature (e.g., Baker, 1984a, 1984b) and previous experiments in this thesis.

Also, children generally made more correct sense judgements for consistent than general knowledge violation error passages. This suggests that children set appropriate standards of evaluation for error detection. However, for the two experiments there was a different pattern of findings for the relation between correct sense judgments for the unrelated general knowledge violation and consistent passages. In Experiment 4, children made more correct sense judgements for the consistent than the unrelated general knowledge violation passages, whereas in Experiment 5 children made a similar number of correct sense judgements for the two passages. Examination of the means for both experiments shows that children were more likely to detect unrelated general knowledge violation errors in Experiment 5 than Experiment 4. However, this difference can be considered spurious because independent samples *t*-tests comparing performance on the unrelated general knowledge violations between age groups for the two experiments revealed that this difference was not significant (*t*s < 1, ns for both age groups).

In addition, in both experiments children made more correct sense judgements for the unrelated than the related general knowledge violation error passages. Thus, suggesting that the explicitness of the error information influenced children's performance on the sense question. It appears that children encountered more difficulty in successfully judging whether or not a passage made sense when it included error information that was closely related to the standard state of affairs. This may be because it is less clear whether or not this information is consistent with world knowledge. So, children may be less likely to identify the error information as problematic and more likely to attempt to resolve the error information.

Next, the reading time findings will be considered. A similar pattern of real-time findings was found in both experiments. This suggests that presentation style did not influence children's real-time reading behaviour. Reading times for the target sentence were longer when it contained a general knowledge violation than when it was consistent. This finding indicates that children detected the general knowledge violation errors. Probably, the extra time children spent on the target error sentence reflects their effort to compare information with their prior world knowledge and consider possible resolutions.

In contrast to the sense question findings, children's reading times for the target sentence were similar for the related and unrelated general knowledge violation passages. This finding suggests that the explicitness of error information did not affect children's real-time reading behaviour. Children slowed down their processing rate to a similar extent for both related and unrelated general knowledge violation information. Thus, it is possible that children experience similar processing demands when encountering related and unrelated general knowledge violation errors. However, here the reader should note a potential limitation of the materials included in experiments in this chapter. The explicitness of error information was decided by the experimenter and verified by experienced adult readers. It may be the case that children's opinion on the explicitness of this information differs from that of adults. Thus, the error information included in the unrelated and related general knowledge violation passages may not be sufficiently different to result in increased processing demands, and consequently differences in reading times.

It is of interest that a similar pattern of real-time findings was found in both the analysis for all responses (and thus regardless of accuracy) and correct-only responses to the sense question. This suggests that even when children failed to correctly answer the sense question at some level they were detecting the error and slowing down their processing rate. Thus, it seems that children generated an internal signal of comprehension failure during reading, but failed to notice the signal or adequately interpret its meaning (e.g., Harris et al., 1981). It is likely that the slight difference between the analysis for all responses and correct-only responses observed in Experiment 4 was due to the analysis for all responses including reading behaviours associated with instances when the sense question was answered incorrectly. Examples include instances when children have not spent sufficient time reading and integrating information to detect an error, or instances when children have spent an extended period resolving erroneous information so that it is no longer considered problematic. It was not possible to analyse incorrect-only reading times because of reduced participant numbers and consequent issues with further reductions in power and statistical sensitivity.

There was no evidence of processing spillover effects in either experiment. The general knowledge violation effects were limited to the reading time for the target sentence and did not extend into the post-target sentence. This finding suggests that children completed processing the error information in the target sentence before moving on to the post-target sentence. Thus, it may be that the general knowledge violation information did not impose sufficient demands on children's information processing resources to influence the time spent reading the post-target sentence. So, children were able to detect the general knowledge violations quickly and easily.

Finally, the relationship between monitoring performance and independent measures of word reading, reading comprehension and working memory will be considered. For the older age group, there were no moderate or strong associations between the monitoring performance and the independent measures. Because children in the older age group performed well on the monitoring task (> 77%), this finding may reflect the possibility that there was limited variance to be explained by the independent measures. For the younger age group, correlations revealed that monitoring performance was related to word and nonword reading, reading accuracy and numerical working memory. However, the regression analyses revealed that word and nonword reading was the only significant predictor of performance on the sense question for the unrelated general knowledge violation passages. Thus, it seems that in younger children proficient comprehension monitoring of general knowledge violations is particularly linked to word reading skill. As mentioned previously, this is perhaps unsurprising as it has been suggested that slow or inaccurate word reading may negatively affect comprehension monitoring by placing too large a demand on children's information processing capacity, leaving little remaining for undertaking important monitoring processes (e.g., Perfetti, 1985).

A potential limitation of the experiments included in this chapter is the relatively small number of participants included in each experiment. Such samples sizes are typical in developmental studies investigating the psychology of language (e.g., Oakhill et al., 2005; van der Schoot et al., 2011). Nonetheless, the issue of statistical sensitivity should be noted; low power resulting from small sample sizes increases the risk that the direction of effect and estimated effect size are inaccurate (e.g., Field, 2005).

In summary, findings revealed developmental differences in correctly judging whether or not a passage made sense. Older children made more correct sense judgments than younger children. In general, children made more correct sense judgements for consistent than general knowledge violation passages. Children also made more correct sense judgements for unrelated than related general knowledge violation passages. This finding suggests that the explicitness of the error information influenced children's performance on the sense question. Both age groups had similar real-time reading behaviour. In general, reading times were longer for the target sentence when it contained a general knowledge violation than when it was consistent. However, in contrast to the sense question findings, reading times were similar for the related and unrelated general knowledge violation passage. This suggests that the explicitness of the error information did not affect children's real-time reading behaviour. In addition, the pattern of findings was similar for both experiments suggesting that the presentation style of passages did not influence children's monitoring of general knowledge violation errors. Further, for young children in particular, word and nonword reading was found to be important in the successful monitoring of general knowledge violation errors.

CHAPTER SIX

INTERNAL INCONSISTENCIES: READING TIME EXPERIMENTS INVESTIGATING THE DISTANCE BETWEEN INCONSISTENT INFORMATION

The main aim of the two experiments reported in this chapter was to investigate whether the distance between inconsistent information influenced children's error detection and real-time reading behaviour. In addition, as in Chapter 5 the presentation style of text was manipulated (each sentence was presented on a separate line, Experiment 6 vs. each sentence was presented in continuous text, Experiment 7) to determine if this influenced children's error detection and real-time reading behaviour.

Introduction

Previous experiments in this thesis demonstrated developmental differences in children's correct responses to the sense question: Older children were better at correctly judging whether or not a passage containing an internal inconsistency error made sense than younger children. In contrast, the reading time measure (used in Experiment 3, see Chapter 4 for a full description) suggested that both age groups had similar real-time reading behaviour: Regardless of age, children took longer to read the target sentence when it contained an internal inconsistency than when it did not. The two experiments reported in this chapter manipulated the distance between inconsistent information to gain a greater insight into whether this aspect of the error influences the likelihood of error detection. Children were provided with instructions that alerted them to the presence of errors in texts. Similar to Experiments 4 and 5, to determine the influence of text presentation on real-time reading behaviour the experiments reported in this chapter contrast two different types of presentation (each sentence was presented on a separate line, Experiment 6 vs. each sentences was presented in continuous text, Experiment 7). In both conditions, text presentation allowed children to look back at previous text. As in Experiments 3-5, children's reading times were recorded so that their real-time reading behaviour could be assessed. To enable comparison with the previous experiments in this thesis, the same age groups (7- to 8-year-olds and 9- to 10-year-olds) participated in each experiment.

Research has shown that the distance between inconsistent information can influence children's error detection. Typically, the greater the distance between the two pieces of inconsistent information, the greater the problems children encounter in monitoring their comprehension for errors in texts (e.g., Ackerman, 1984b; Oakhill et al., 2005; van der Schoot et al., 2011; Zabrucky & Ratner, 1986). It has been suggested that such problems derive from the working memory demands involved in the integration of the two pieces of inconsistent information (e.g., Albrecht & O'Brien, 1993). For internal inconsistency errors to be detected, the two pieces of inconsistent information must be concurrently active in working memory. The greater the distance between the inconsistent information, the greater the storage and processing demands involved in ensuring that the two pieces of information are co-activated in working memory. To extend our understanding of the factors that influence children's monitoring of internal inconsistency errors, the distance between inconsistent information was manipulated.

Further, there were two key methodological differences between internal inconsistency passages used in current experiments and those used in previous experiments. First, the length of the passages was increased and second, matched versions of each passage manipulation were created. These methodological differences are the same as those stated for the general knowledge violation passages used in Chapter 5 (see the Introduction of Chapter 5 for a more detailed account of these differences). Note here that in addition to allowing spillover effects to be investigated, increasing the length of passages enabled the distance between inconsistent information to be manipulated.

6.1. Experiment 6: Comprehension monitoring of near and far internal inconsistencies when successive sentences in a passage are presented on separate lines

In line with previous research (e.g., Baker, 1984a) and the findings from previous experiments, it was predicted that in comparison to older children, younger children would make fewer correct responses to the question ‘did the story make sense’. It was also predicted that children would make more correct responses to the sense question for the near internal inconsistency passages than for the far internal inconsistency passages. In line with previous research (e.g., Albrecht & O’Brien, 1993; van der Schoot et al., 2011), it was predicted that reading times would be longer for the target sentence when it contained an error (i.e., an internal inconsistency) than when it was consistent. As stated previously, this pattern would indicate detection of errors in real-time during reading and suggest that comprehension monitoring has taken place. In addition, it was predicted that reading times would be longer for the target sentence when it contained a far than a near internal inconsistency. The greater the distance between inconsistent information, the greater the storage and processing demands involved in ensuring that inconsistent information is concurrently active in working memory. Thus, it is possible that children will need to slow down their processing rate to integrate information from successive sentences into their expanding situation model and evaluate the consistency of information.

Method

In Experiment 6, the participants (see Participants Section below for exceptions), word reading, reading comprehension and working memory measures and procedure were the same as in Experiment 4. The only differences between the two experiments were the comprehension monitoring materials and the list construction of these materials. Here, only these differences are described along with a reminder of the method of stimulus presentation.

Participants

One participant from the 7- to 8-year-old age group and two participants from the 9- to 10-year-old age group in Experiment 4 did not participate in the current experiment. In addition, two new participants were included in the younger age group and one new participant was included in the older age group. Thus, 26 children (12 boys, 14 girls) with a mean age of 7;9 years (*SD* = 4 months, range = 7;3 – 8;3 years) and 28 children (10 boys, 18 girls) with a mean age of 9;8 years (*SD* = 4 months, range = 9;3 – 10;4 years) participated.

Materials

Sixteen passages from Experiments 1 and 3 were modified (breakdown: 2 passages from Experiment 1 (used a practice passages), and 14 passages from Experiment 3), and five new passages were created for the purposes of this experiment. Thus, there were a total of 21 eight-sentence passages: 18 experimental passages and three practice passages. All passages were a similar length (*M* = 79.31, *SD* = 3.31, range = 75 – 85) and had a similar Flesch-Kincaid Grade readability (*M* = 2.63, *SD* = .32, range = 2.09 – 3.18). For the experimental passages, there were three versions of each passage: near internal inconsistency, far internal inconsistency and consistent (see Appendix 4.1 for a full list of passages). In the near internal inconsistency version, contradictory information appeared in sentences two and four. In the far internal inconsistency version, contradictory information appeared in sentences two and six. The near and far internal inconsistency versions were constructed such that the position of the inconsistent sentences could be changed without changing the wording of any sentence in a passage. In the consistent version, the inconsistent information in sentences two and six (far internal inconsistency version) was resolved. This required only minor changes to the wording of either sentence two or six. An example of a passage with all versions is shown in Table 6.1.

Table 6.1. Example of a passage

Josh has a rabbit named Bugs.
<u>Bugs lives in a cage in the living room, and never goes outside.</u>
(Bugs lives in a cage in the living room, but likes going outside.)
Josh's Mum and Dad bought Bugs as a birthday present for Josh.
*
Bugs has grey and white fur that is very soft.
Josh feeds Bugs rabbit food and water.
<u>*Everyday, Bugs plays in the garden on the grass.</u>
As a special treat, Josh gives Bugs a carrot to eat every Friday.
Josh really likes having Bugs as his pet.

Note. Far internal inconsistency version presented. Near internal inconsistency version indicated by the asterisk and consistent version shown in parenthesis.

A pilot study with 18 adults confirmed that all errors were identifiable. Any unintended errors that were pointed out by the adult readers were modified before the experiment was conducted with children.

Self-paced moving window method

As in Experiment 4, reading times were collected by means of the self-paced sentence-by-sentence moving window method (e.g., Just et al., 1982). In this method, participants are presented with passages of text on a computer screen in such a way that each sentence is masked by an 'x' and sentences are presented one at a time. When the participant has finished comprehending a sentence, they press a button to move on to the next sentence. In the current experiment, once a sentence had been presented to the participant it remained on the screen so that children could go back and re-read previous sentences of a passage. Sentence reading times were defined as beginning when the sentence was first revealed and lasting until the next button press.

Procedure

As in Experiment 4, passages were assembled into three lists. Each list consisted of 18 experimental passages: six near internal inconsistency, six far internal inconsistency and six consistent passages. Each list also had three practice passages: one near internal inconsistency, one far internal inconsistency and one consistent passage. The same practice passages were used for each list. The different versions of each experimental passage were counterbalanced across the lists so that each version of a given passage occurred only once in each list. Following the practice, the first three passages of each list were arranged in a constrained order such that a near internal inconsistency passage was encountered first, a far internal inconsistency passage second and a consistent passage third. The remaining passages were presented in a pseudorandomised order in which no more than two of the same type occurred consecutively within a list. Each participant saw one list.

Results

Data filtering

Reading times per syllable (in milliseconds) were calculated for each sentence. Data filtering and preparation was the same as that used in Chapter 5 (see the Results Section of Experiment 4 for further details). This data filtering eliminated approximately 2.5% of the data analysed. No reading times were more than 2.5*SD* from the mean. Before analysis the Q-Q plots for all measures were examined for each age group separately. The plots indicated that the distribution of the data was acceptable.

Correct sense judgements

The mean number correct for the near internal inconsistency, far internal inconsistency and consistent passage types was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The mean numbers of correct sense judgements made for each passage type by age group are shown in Table 6.2.

Table 6.2. Mean numbers (with standard deviations) of correct judgements by age group and passage type

	Age group		
Passage type	**7- to 8-year-olds (n = 26)**	**9- to 10-year-olds (n = 28)**	**Total**
Near internal inconsistency (maximum = 6)	3.83 (1.56)	4.86 (1.11)	4.35 (1.34)
Far internal inconsistency (maximum = 6)	3.88 (1.68)	4.89 (1.17)	4.39 (1.43)
Consistent (maximum = 6)	5.19 (.80)	5.04 (1.07)	5.13 (.94)
Total (maximum = 6)	4.30 (1.35)	4.93 (1.12)	

There was a significant main effect of age group ($F(1,52) = 6.72$, $p = .01$, $\eta p^2 = .11$) because the older children made more correct sense judgements than the younger children. There was also a main effect of passage type ($F(2,104) = 8.38$, $p < .001$, $\eta p^2 = .14$). Children made more correct sense judgements for the consistent passages than both the near internal inconsistency ($t(53) = 3.24$, $p = .01$, $d = .67$) and the far internal inconsistency passages ($t(53) = 2.98$, $p = .01$, $d = .60$). In addition, children made a similar number of correct sense judgements for the near and far internal inconsistency passages ($t(53) < 1$, ns).

There was a significant interaction between age group and passage type ($F(2,104) = 5.10$, $p = .01$, $\eta p^2 = .09$) which is shown in Figure 6.1. The interaction arose because older children made a similar number of correct sense judgements for all passage types (all $ts(27) < 1$, ns). Whereas younger children made more correct sense judgements for the consistent than the near internal inconsistency ($t(25) = 4.44$, $p < .001$, $d = 1.10$) and far internal inconsistency ($t(25) = 3.68$, $p = .01$, $d = 1.00$) passages.

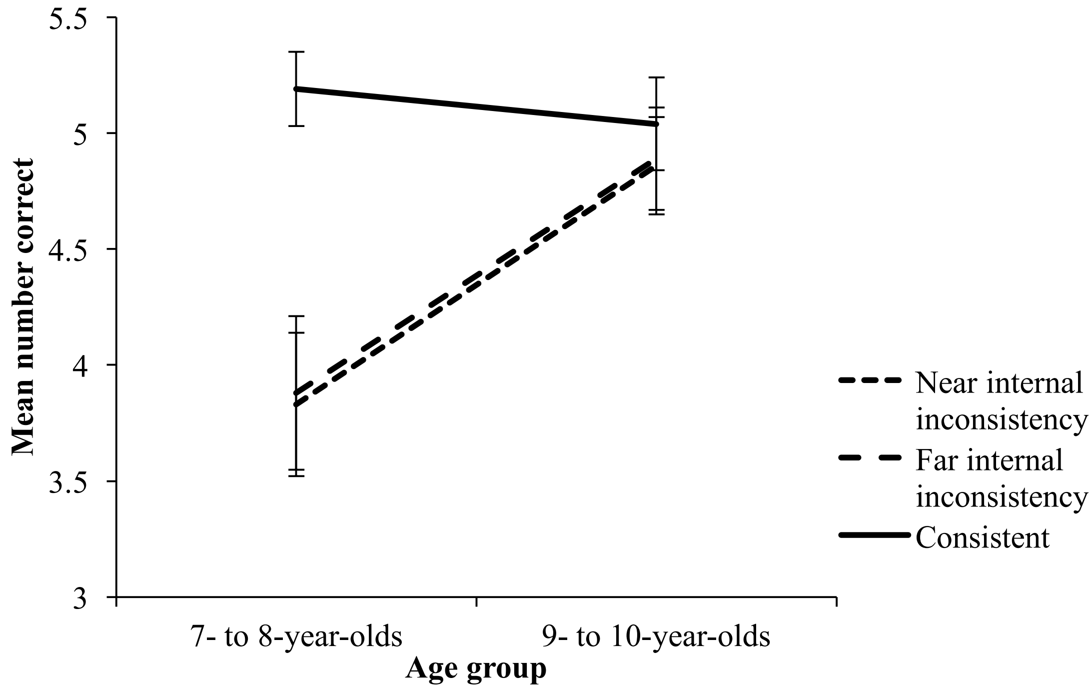

Figure 6.1. Mean numbers (+/- standard error) of correct judgements by age group and passage type

Reading times

Two analyses comparing the target sentence for the near internal inconsistency, far internal inconsistency and consistent passages were conducted. The first includes all responses (correct and incorrect) to the sense question; the second includes correct-only responses to the sense question.

Analysis of reading times for the target sentence for all responses (correct and incorrect) to the sense question. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 6.3.

Table 6.3. Mean reading times (ms/syllable) (with standard deviations) for the target sentence and post-target sentence by age group, passage type and question accuracy

		Age group			
		7- to 8-year-olds		**9- to 10-year-olds**	
Passage type	**Question accuracy**	**Target sentence**	**Post-target sentence**	**Target sentence**	**Post-target sentence**
Near internal inconsistency	**Overall**	513.53 (168.91)	452.58 (166.43)	357.14 (115.23)	275.75 (83.90)
	Correct-only	524.25 (174.25)[a]	447.73 (205.32)[a]	361.87 (111.83)	264.45 (84.32)
Far internal inconsistency	**Overall**	487.33 (148.86)	454.15 (166.48)	371.44 (94.69)	305.87 (111.43)
	Correct-only	506.77 (149.30)[a]	473.31 (268.04)[a]	381.11 (96.96)	297.64 (109.62)
Consistent	**Overall**	491.72 (164.17)	453.58 (157.08)	334.90 (93.53)	343.27 (99.59)
	Correct-only	464.83 (158.87)[a]	451.00 (166.18)[a]	335.73 (86.24)	333.36 (97.71)

Note. Overall refers to both correct and incorrect responses. Unless indicated in the table by superscript, n = 26 for 7- to 8-year-olds and n = 28 for 9- to 10-year-olds. [a] n = 23. Missing data is due to one child having no data for correct responses for near internal inconsistency passages and two children having no data for correct responses for far internal inconsistency passages. Note missing data is for three different children.

There was a significant main effect of age group ($F(1,52) = 19.48$, $p < .001$, $\eta p^2 = .27$) because the older children ($M = 354.49$, $SD = 101.15$) read the target sentence faster than the younger children ($M = 497.53$, $SD = 160.65$). However, there was no main effect of passage type ($F(2,104) = 1.28$, $p = .28$, $\eta p^2 = .02$). Children had similar reading times for the target sentence when it contained a near internal inconsistency ($M = 435.34$, $SD = 142.07$), far internal inconsistency ($M = 429.39$, $SD = 121.78$) and when it was consistent ($M = 413.31$, $SD = 128.85$). The interaction between age group and passage type was not significant ($F(2,104) = 1.36$, $p = .26$, $\eta p^2 = .03$).

Analysis of reading times for the target sentence for correct-only responses to the sense question. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 6.3.

Similar to the previous analysis there was a significant main effect of age group ($F(1,49) = 18.81$, $p < .001$, $\eta p^2 = .28$) because the older children read the target sentence faster ($M = 359.57$, $SD = 98.34$) than the younger children ($M = 498.62$, $SD = 160.81$). In contrast to the previous analysis, there was also a significant main effect of passage type ($F(2,98) = 5.15$, $p = .01$, $\eta p^2 = .10$). Children had longer reading times for the target sentence when it contained a near internal inconsistency ($M = 443.06$, $SD = 143.04$) ($t(52) = 1.96$, $p = .06$, $d = .32$) and a far internal inconsistency ($M = 443.94$, $SD = 123.13$) ($t(51) = 2.95$, $p = .01$, $d = .36$) than when it was consistent. In addition, children had similar reading times for the target sentence when it contained a near and a far internal inconsistency ($t(50) < 1$, ns). As in the previous analysis, the interaction between age group and passage type was not significant ($F(2,98) < 1$, ns).

Analysis of spillover effects. Reading times for the post-target sentence for all responses (correct and incorrect) and correct-only responses to the sense question were analysed to explore possible spillover effects. As the near internal inconsistency errors were inserted into sentence four and the far internal inconsistency errors into sentence six, the post-target sentence was sentence five or seven, depending on the passage type.

All of reading times for the post-target sentence for all responses (correct and incorrect) to the sense question. The mean reading time for the post-target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 6.3.

There was a significant main effect of age group ($F(1,52) = 19.13$, $p < .001$, $\eta p^2 = .27$) because the older children read the post-target sentence faster ($M = 354.59$, $SD = 101.15$) than the younger children ($M = 483.77$, $SD = 163.33$). There was also a significant main effect of passage type ($F(2,104) = 3.45$, $p = .04$, $\eta p^2 = .06$). Children had longer reading times for the post-target sentence when it contained a near internal inconsistency ($M = 364.17$, $SD = 125.17$) than when it was consistent ($M = 398.43$, $SD = 128.34$). In addition, children had similar reading times for the post-target sentence when it contained a far internal inconsistency ($M = 380.01$, $SD = 138.96$) and when it was consistent, and when it contained a near and far internal inconsistency.

There effect of age group was qualified by a significant interaction with passage type ($F(2,104) = 3.26$, $p = .04$, $\eta p^2 = .06$), which is shown in figure 6.2. The interaction arose because younger children showed no effect of passage type (all $ts(25) < 1$, *ns*), whereas older children showed an effect of passage type. Reading times for the post-target sentence were longer for the consistent passages than the near internal inconsistency ($t(27) = 5.04$, $p < .001$, $d = .73$) and far internal inconsistency ($t(27) = 2.17$, $p = .04$, $d = .35$) passages.

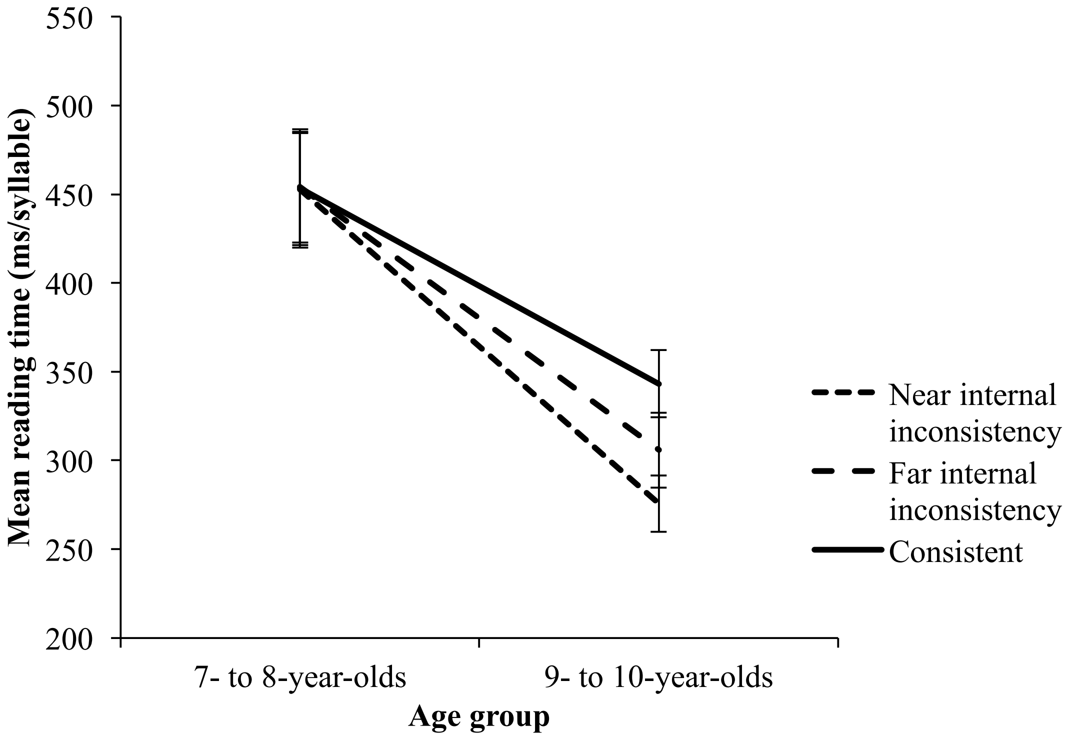

Figure 6.2. Mean reading times (ms/syllable) (+/- standard error) for the post-target sentence by age group and passage type

Analysis of reading times for the post-target sentence for correct-only responses to the sense question. The mean reading time for the post-target sentence was the dependent variable in a 2 (age group) x 3 (passage type) ANOVA. The means for the near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 6.3.

Similar to the previous analysis, there was a significant main effect of age group ($F(1,49) = 15.89$, $p < .001$, $\eta p^2 = .25$) because the older children read the post-target sentence faster ($M = 298.48$, $SD = 97.23$) than the younger children ($M = 457.35$, $SD = 213.18$). In contrast to the previous analysis, there was no significant main effect of passage type ($F(2,98) = 1.92$, $p = .15$, $\eta p^2 = .04$), and the interaction between age group and passage type was not significant ($F(2,98) = 1.68$, $p = .19$, $\eta p^2 = .03$).

Summary

The main aim of Experiment 6 was to investigate how the distance between inconsistent information influenced 7- to 8-year-old and 9- to 10-year-old children's comprehension monitoring of internal inconsistency errors when successive sentences in a passage were presented on separate lines. As predicted, and in line with previous experiments in this thesis, differences between age groups were found: Older children made significantly more correct sense judgements than younger children. In an extension of previous experiments, the distance between inconsistent information was manipulated. These novel findings are outlined below.

Children made more correct sense judgements for consistent than both the near and far internal inconsistency error passages. However, in contrast to predictions, children made a similar number of correct sense judgements for the near and far internal inconsistency passages. There was an interaction between age group and passage type, which revealed that older children made a similar number of correct sense judgements for all passage types. Whereas younger children made more correct sense judgements for the consistent than the near and far internal inconsistency passages.

The reading time findings revealed different patterns for all responses and correct-only responses to the sense question. Only in the analysis for correct-only responses was a main effect of passage type found for the target sentence: Children had longer reading times for the target sentence when it contained a near and far internal consistency than when it was consistent. In addition, the analysis for correct-only responses revealed that in contrast to predictions, children's reading times for the target sentence were similar when it contained a near and far internal inconsistency.

Further, only in the analysis for all responses to the sense question was a main effect of passage type found for the post-target sentence: Younger children's reading times for the post-target sentence were similar for all passage types (near internal inconsistency, far internal inconsistency and consistent). In contrast, older children had longer reading times for the post-target sentence when it was consistent than when it contained either a near or far internal inconsistency. However, older children's reading times for the post-target sentence were similar when it contained a near and far internal inconsistency.

As outlined in the Summary for Experiment 4, it is possible that these age group and error findings are due to the separate line presentation style of the passages. Thus, Experiment 7 was conducted to investigate how the distance between inconsistent information influenced children's comprehension monitoring of internal inconsistency errors when successive sentences in a passage were presented in connected text.

6.2. Experiment 7: Comprehension monitoring of adjacent, near and far internal inconsistencies when successive sentences in a passage are presented in connected text

In Experiment 6, error information was placed in sentences two and four, and sentences two and six creating the near and far manipulations, respectively. In Experiment 7, a third distance manipulation, the adjacent internal inconsistency, was created. This was to allow investigation of the influence of having no filler text intervening between inconsistent information on children's comprehension monitoring, and allow this to be compared to the near and far internal inconsistency manipulations. Passages were presented in connected text. As in Experiment 6, main effects of age group and passage type were predicted.

Method

In Experiment 7, the participants (see Participants Section below for exceptions), word reading, reading comprehension and working memory measures and procedure were the same as in Experiment 5. The only differences between the two experiments were the comprehension monitoring materials and the list construction of materials. Here, only these differences are described.

Participants

One participant from the 7- to 8-year-old and 9- to 10-year-old age groups in Experiment 5 did not participate in the current experiment. Thus, 21 children (9 boys, 12 girls) with a mean age of 8;2 years (*SD* = 4 months, range = 7;9 – 8;8 years) and 22 children (12 boys, 10 girls) with a mean age of 10;4 years (*SD* = 3 months, range = 9;10 – 10;9 years) participated.

Materials

Passages from Experiment 6 underwent minor revisions to address issues raised by children, accommodate the adjacent manipulation and the new presentation style. Six new passages were also created for the purposes of this experiment. Thus, there were a total of 27 passages: 24 experimental passages and three practice passages. All passages were a similar length (*M* = 81.69, *SD* = 3.77, range = 75 – 88) and had a similar Flesch-Kincaid Grade readability (*M* = 2.79, *SD* = .40, range = 1.60 – 3.40). There were four versions of each experimental passage: adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent (see Table 6.4 for an example and Appendix 4.2 for a full list of passages). In the adjacent internal inconsistency version, conflicting information appeared in sentences two and three. The near internal inconsistency, far internal inconsistency and consistent versions remained the same as in Experiment 6. As in previous experiments, for the new passages a pilot study with 18 adults confirmed that all errors were identifiable.

Table 6.4. Example of a passage

Josh has a rabbit named Bugs.
Bugs lives in a cage in the living room, and never goes outside.
(Bugs lives in a cage in the living room, but likes going outside.)
*
Josh's Mum and Dad bought Bugs as a birthday present for Josh.
*
Bugs has grey and white fur that is very soft.
Josh feeds Bugs rabbit food and water.
*Everyday, Bugs plays in the garden on the grass.
As a special treat, Josh gives Bugs a carrot to eat every Friday.
Josh really likes having Bugs as his pet.

Note. Far internal inconsistency version presented. Adjacent internal inconsistency version indicated by the first asterisk, near internal inconsistency version indicated by the second asterisk and consistent version shown in parenthesis.

Procedure

Passages were assembled into four lists. Each list consisted of 24 experimental passages: six adjacent internal inconsistency, six near internal inconsistency, six far internal inconsistency and six consistent passages. Each list also had three practice passages: one adjacent internal inconsistency, one far internal inconsistency and one consistent passage. The same practice passages were used for each list. The different versions of each experimental passage were counterbalanced across the lists so that each version of a given passage occurred only once in each list. Passages were presented in a pseudorandomised order in which no more than two of the same type occurred consecutively within a list. Each participant saw one list.

As in Experiment 5, each passage was displayed in its entirety on one computer screen, one sentence at a time. Each sentence was presented in its naturally occurring position from left to right on each line of the screen, with successive lines appearing below each other in connected text. In line with the self-paced moving window method, until presented each character (with the exception of spaces and punctuation) of the text was masked by an 'x'.

Results

Reading times per syllable (in milliseconds) were calculated for each sentence. Data filtering and preparation was the same as the previous experiment. This data filtering eliminated approximately 3% of the data analysed. No reading times were more than 2.5*SD* from the mean. Before analysis the Q-Q plots for all measures were examined for each age group separately. The plots indicated that the distribution of the data was acceptable.

Correct sense judgements

The mean number correct for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passage types was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The mean numbers of correct sense judgements made for each passage type by age group are shown in Table 6.5.

Table 6.5. Mean numbers (with standard deviations) of correct judgements by age group and passage type

	Age group		
Passage type	**7- to 8-year-olds (n = 21)**	**9- to 10-year-olds (n = 22)**	**Total**
Adjacent internal inconsistency (maximum = 6)	4.19 (.97)	5.07 (1.07)	4.63 (1.02)
Near internal inconsistency (maximum = 6)	4.22 (1.22)	4.27 (1.17)	4.25 (1.20)
Far internal inconsistency (maximum = 6)	3.62 (1.48)	4.21 (1.19)	3.92 (1.34)
Consistent (maximum = 6)	4.76 (1.04)	5.20 (.76)	4.98 (.90)
Total (maximum = 6)	4.20 (1.18)	4.69 (1.05)	

The pattern of results was similar to that found for the previous experiment. As before, there was a significant main effect of age group ($F(1,41) = 4.50$, $p = .04$, $\eta p^2 = .10$) because the older children made more correct sense judgements than the younger children. As before, there was also a main effect of passage type ($F(3,123) = 9.78$, $p < .001$, $\eta p^2 = .19$). Children made more correct sense judgements for the consistent passages than the near internal inconsistency ($t(42) = 3.66$, $p = .01$, $d = .69$) and far internal inconsistency ($t(42) = 5.35$, $p < .001$, $d = .93$) passages. Children also made a similar number of correct sense judgements for the near and far internal inconsistency passages ($t(42) = 1.40$, $p = .17$, $d = .26$). The addition of the adjacent internal inconsistency manipulation yielded novel results. Children made a similar number of correct sense judgements for the adjacent and consistent passages ($t(42) = 1.77$, $p = .09$, $d = .36$), and adjacent and near internal inconsistent passages ($t(42) = 1.80$, $p = .08$, $d = .34$). In addition, children made more correct sense judgements for the adjacent internal inconsistency passages than the far internal inconsistency passages ($t(42) = 3.41$, $p = .01$, $d = .60$). In contrast to Experiment 6, the interaction between age group and passage type was not significant ($F(3,123) = 1.38$, $p = .25$, $\eta p^2 = .03$).

Reading times

As in the previous experiment, two analyses comparing the target sentence for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages are reported. The first includes all responses (correct and incorrect) to the sense question, the second includes correct-only responses to the sense question.

Analysis of reading times for the target sentence for all responses (correct and incorrect) to the sense question. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The means for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 6.6.

The pattern of results was different to that found for the previous experiment. As before, there was a significant main effect of age group ($F(1,41) = 4.65$, $p = .04$, $\eta p^2 = .10$) because the older children ($M = 394.85$, $SD = 122.83$) read the target sentence faster than the younger children ($M = 472.65$, $SD = 140.35$). In contrast to the previous experiment, there was also a main effect of passage type ($F(3,123) = 6.32$ $p = .01$, $\eta p^2 = .13$). Children had longer reading times for the target sentence when it contained an adjacent internal inconsistency ($M = 460.57$, $SD = 147.19$), near internal inconsistency ($M = 435.16$, $SD = 120.11$) and far internal inconsistency ($M = 441.40$, $SD = 140.91$) than when it was consistent ($M = 397.87$, $SD = 118.16$) ($t(42) = 4.11$, $p < .001$, $d = .47$ for adjacent internal inconsistency and consistent passages; $t(42) = 2.78$, $p = .01$, $d = .31$ for near internal inconsistency and consistent passages; $t(42) = 2.90$, $p = .01$, $d = .33$ for far internal inconsistency and consistent passages). In addition, children had similar reading times for the target sentence when it contained an adjacent, near or far internal inconsistency (all $ts(42) < 1.82$, $ps > .08$). As in Experiment 6, the interaction between age group and passage type was not significant ($F(3,123) = 2.21$, $p = .09$, $\eta p^2 = .05$).

Analysis of reading times for the target sentence for correct-only responses to the sense question. The mean reading time for the target sentence was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The means for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 6.6.

Table 6.6. Mean reading times (ms/syllable) (with standard deviations) for the target sentence and post-target sentence by age group, passage type and question accuracy

		Age group			
		7- to 8-year-olds (n = 21)		9- to 10-year-olds (n = 22)	
Passage type	**Question accuracy**	**Target sentence**	**Post-target sentence**	**Target sentence**	**Post-target sentence**
Adjacent internal inconsistency	**Overall**	522.24 (159.99)	432.13 (128.90)	398.90 (134.38)	312.97 (121.19)
	Correct-only	509.21 (150.22)	428.77 (137.39)	401.95 (142.19)	293.35 (133.81)
Near internal inconsistency	**Overall**	465.50 (131.21)	417.84 (119.88)	404.82 (109.00)	361.40 (113.67)
	Correct-only	473.97 (127.71)	416.66 (116.02)	410.38 (120.26)	361.27 (130.52)
Far internal inconsistency	**Overall**	469.40 (146.19)	418.70 (142.42)	413.39 (135.62)	363.57 (140.80)
	Correct-only	492.83 (174.17)	423.41 (187.42)	409.46 (123.95)	373.81 (145.91)
Consistent	**Overall**	433.46 (124.00)	448.91 (148.36)	362.28 (112.32)	357.46 (140.93)
	Correct-only	422.38 (131.16)	434.92 (156.70)	363.49 (118.96)	356.61 (143.35)

Note. Overall refers to both correct and incorrect responses.

The pattern of results was similar to that found for the previous experiment, and the same as that found for the analysis above for all (correct and incorrect) responses to the sense question. As before, there was a significant main effect of age group ($F(1,41) = 4.98$, $p = .03$, $\eta p^2 = .11$) because the older children read the target sentence faster ($M = 396.32$, $SD = 126.34$) than the younger children ($M = 474.60$, $SD = 145.82$). As before, there was also a significant main effect of passage type ($F(3,123) = 4.85$, $p = .01$, $\eta p^2 = .11$). However, the pattern of difference was not the same as that found in Experiment 6. Children had longer reading times for the target sentence when it contained an adjacent internal inconsistency ($M = 455.58$, $SD = 146.21$), near internal inconsistency ($M = 442.18$, $SD = 123.99$) and far internal inconsistency ($M = 451.15$, $SD = 149.06$) than when it was consistent ($M = 392.94$, $SD = 125.06$) ($t(42) = 3.27$, $p = .01$, $d = .46$ for adjacent internal inconsistency and consistent passages; $t(42) = 3.07$, $p = .01$, $d = .40$ for near internal inconsistency and consistent passages; $t(42) = 3.34$, $p = .01$, $d = .42$ for far internal inconsistency and consistent passages). In addition, children had similar reading times for the target sentence when it contained an adjacent, near or far internal inconsistency (all $ts(42) < 1$, ns). As before, the interaction between age group and passage type was not significant ($F(3,123) < 1$, ns).

Analysis of spillover effects. Separate analyses for the post-target sentence for all responses (correct and incorrect) and correct-only responses to the sense question were conducted. The mean reading time for the post-target sentence was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The means for adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 6.6.

The pattern of results was the same as that found in the previous experiment for correct-only responses to the sense question. There was a significant main effect of age group ($F(1,41) = 5.21$, $p = .03$, $\eta p^2 = .11$; $F(1,41) = 4.49$, $p = .04$, $\eta p^2 = .10$, for all and correct-only responses analyses, respectively) because the older children read the post-target sentence faster than the younger children. There was no significant main effect of passage type (F(3,123) = 1.23, p = .30, ηp^2 = .03; F(3,123) = 1.63, p = .19, ηp^2 = .04). Further, the interaction between age group and passage type was not significant (F(3,123) = 1.83, p = .15, ηp^2 = .04; F(3,123) = 2.11, p = .10, ηp^2 = .05).

Summary

The main aim of Experiment 7 was to investigate how the distance between inconsistent information influenced 7- to 8-year-old and 9- to 10-year-old children's comprehension monitoring of internal inconsistency errors when successive sentences in a passage were presented in connected text. As predicted, and in line with the findings from Experiment 6, differences in age groups were found: Older children made significantly more correct sense judgements than younger children. Novel findings are outlined below.

Similar to the findings for Experiment 6, children made more correct sense judgements for consistent than near and far internal inconsistency error passages, and a similar number of correct sense judgements for near and far internal inconsistency passages. The addition of the adjacent internal inconsistency manipulation provided additional information. Children made more correct sense judgements for adjacent than far internal inconsistency passages, and a similar number of correct sense judgements for adjacent internal inconsistency and consistent, and adjacent and near internal inconsistency passages.

The reading time findings were different to those for Experiment 6. The pattern of findings was the same for both the analysis for all responses (and thus regardless of accuracy) and for correct-only responses to the sense question. Children had longer reading times for the target sentence when it contained an internal inconsistency than when it was consistent. In contrast to predictions, children's reading times for the target sentence were similar when it contained an adjacent, near or far internal inconsistency. In addition, children's reading times for the post-target sentence were similar for all passage types (adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent).

6.3. The relationship between comprehension monitoring and word reading, reading comprehension and working memory

The relationship between comprehension monitoring and word reading, reading comprehension and working memory were explored. Because children in Experiments 6 and 7 completed the same measures, and sense judgement analyses revealed essentially the same main effects, for reasons of power data from the two experiments were combined for the analyses that follow.

Descriptive statistics

Table 6.7 summarises the mean performance of the two age groups on independent measures of word reading, reading comprehension and working memory for Experiments 6 and 7. Note that the performance of both age groups on the word reading and reading comprehension measures was equal to or above a standard score of 85.

Table 6.7. Mean raw scores, ability scores and standard scores (with standard deviations and t-tests) for word reading, reading comprehension and working memory measures by age group

Measure	7- to 8-year-olds (n = 47)	9- to 10-year-olds (n = 50)	*t*(95)	*d*
Word reading				
Raw score (maximum = 104)	63.43 (7.73)	71.24 (6.98)	5.23***	1.06
Standardised score	116.11 (8.96)	108.34 (10.14)	3.99***	.81
Nonword reading				
Raw score (maximum = 63)	35.34 (9.67)	41.66 (7.71)	3.57**	.72
Standardised score	116.66 (11.11)	112.98 (11.02)	1.64	.33
Reading rate				
Raw score (seconds)	102.40 (27.69)	108.19 (21.94)	1.14	.23
Ability score	67.34 (9.27)	78.58 (7.73)	6.40***	1.32
Standardised score	111.94 (9.89)	109.72 (11.15)	1.36	.21
Word reading accuracy				
Raw score	4.27 (2.41)	3.42 (2.13)	1.83	.37
Ability score	54.45 (8.21)	63.66 (7.55)	5.76***	1.17
Standardised score	110.91 (10.69)	110.18 (10.91)	< 1	-
Reading comprehension				
Raw score (maximum = 8)	4.81 (1.17)	5.18 (1.11)	1.60	.32
Ability score	56.21 (7.23)	63.40 (7.14)	4.93***	1.00
Standardised score	107.26 (9.00)	106.44 (10.66)	< 1	-
Verbal working memory				
Number of items (maximum = 42)	12.85 (5.50)	18.26 (6.63)	4.36***	.89
Number of trials (maximum = 12)	3.81 (1.50)	5.12 (1.52)	4.28***	.87
Span (maximum = 5)	2.23 (.60)	2.68 (.62)	3.60**	.74
Numerical working memory				
Number of items (maximum = 27)	13.32 (5.70)	18.82 (5.38)	4.89***	.99
Number of trials (maximum = 9)	4.19 (1.81)	5.66 (1.67)	4.15***	.84
Span (maximum = 4)	2.32 (.63)	2.96 (.67)	4.85***	.98

Note. Raw scores are presented on the upper row, and where appropriate ability scores on the middle row and standardised scores on the lower row. *$p < .05$, **$p < .01$, ***$p < .001$.

Interrelations between measures

The same correlation analyses were conducted as in Chapter 5. Here the correlation between the TOWRE Sight Word reading and TOWRE Phonemic Decoding measures was *rs* = .82 and .67, *ps* < .001, for older and younger age groups, respectively. Note, despite the moderate association between verbal and numerical working memory measures in both age groups (see Tables 6.8 and 6.9 for values), a composite measure was not created so that the independent relations for each measure could be explored. Measures of monitoring performance were accuracy to the sense question for the adjacent (Experiment 7 only), near and far internal inconsistency passages. Consistent passages were not included in analysis because performance was close to ceiling for both age groups.

For brevity the reader is directed to Tables 6.8 and 6.9 which summarise the correlations for each age group, and only key findings are discussed here. As noted in Chapter 5, many correlations did not reach statistical significance and so will be interpreted in relation to effect sizes. For the 7- to 8-year-old children, there was a moderate association between correct responses to the sense question for the adjacent internal inconsistencies and the reading comprehension measure. For the younger age group, there were also moderate associations between correct responses to the sense question for the near internal inconsistencies and the reading accuracy and reading comprehension measures. The same pattern of associations with the independent measures was also found for correct responses to the sense question for far internal inconsistencies.

For the 9- to 10-year-old children, there were no moderate or strong associations between correct responses to the sense question for adjacent internal inconsistency passages and the other measures. However, for the older age group there were moderate associations between correct responses to the sense question for the near internal inconsistencies and the reading comprehension, verbal working memory and numerical working memory measures. For the older age group, there was also a moderate association between correct responses to the sense question for the far internal inconsistencies and the numerical working memory measure.

Table 6.8. Correlations between measures by age group: 7- to 8-year-olds

	1	2	3	4	5	6	7	8
1. Word and nonword reading	-	.69***	.04	.33*	.24	.20	.18	.09
2. Word reading accuracy	.69***	-	.44**	.39**	.25	.27	.34*	.32*
3. Reading comprehension	.04	.44**	-	.21	.18	.46*	.37*	.31*
4. Verbal working memory	.33*	.39**	.21	-	.46**	.19	.16	.18
5. Numerical working memory	.24	.25	.18	.46**	-	.13	.19	.07
6. Sense question: Adjacent internal inconsistency	.20	.27	.46*	.19	.13	-	.22	.22
7. Sense question: Near internal inconsistency	.18	.34*	.37*	.16	.19	.22	-	.55***
8. Sense question: Far internal inconsistency	.09	.32*	.31*	.18	.07	.22	.55***	-

Note. For the adjacent sense question measure n = 21. For near and far sense question measures n = 47. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 6.9. Correlations between measures by age group: 9- to 10-year-olds

	1	2	3	4	5	6	7	8
1. Word and nonword reading	-	.57***	.31*	.28*	.43**	-.03	.28$^{(*)}$	.08
2. Word reading accuracy	.57***	-	.43**	.35*	.35*	-.23	.16	.13
3. Reading comprehension	.31*	.43**	-	.21	.22	-.39	.31*	.03
4. Verbal working memory	.28*	.35*	.21	-	.44**	-.45	.30*	.11
5. Numerical working memory	.43**	.35*	.22	.44**	-	.23	.39**	.32
6. Sense question: Adjacent internal inconsistency	-.03	-.23	-.39	-.45	.23	-	.23	.47*
7. Sense question: Near internal inconsistency	.28$^{(*)}$	.16	.31*	.30*	.39**	.23	-	.26
8. Sense question: Far internal inconsistency	.08	.13	.03	.11	.32*	.47*	.26	-

Note. For the adjacent sense question measure n = 22. For near and far sense question measures n = 50. $^{(*)}p = .05$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

The role of word reading, reading comprehension and working memory in comprehension monitoring ability

As in Chapter 5, forward stepwise regressions were conducted to explore the specific role of word reading, reading comprehension and working memory in comprehension monitoring ability. The dependent variables were performance on the sense question for adjacent, near and far internal inconsistency passages. Because there were no significant moderate or strong associations between the adjacent internal inconsistency passages and the other measures in the 9- to 10-year-old children, a regression was not conducted for this measure in the older age group. As before, the predictors entered were: word and nonword reading, reading comprehension, verbal working memory and numerical working memory.

For the 7- to 8-year-old children, reading comprehension was the only significant predictor of performance on the sense question for the adjacent, near and far internal inconsistency passages (models explained 22%, 14% and 9% of variance, $\beta = .46$, $p = .03$; $\beta = .37$, $p = .01$; $\beta = .31$, $p = .04$ for adjacent, near and far internal inconsistencies, respectively). In contrast, for the 9- to 10-year-old children numerical working memory was the only significant predictor of performance on the sense question for the near and far internal inconsistency passages (models explained 15% and 10% of variance, $\beta = .39$, $p = .01$; $\beta = .32$, $p = .03$ for near and far internal inconsistencies, respectively).

6.4. Discussion

First, findings for the sense question will be considered. A similar pattern of off-line findings for near and far internal inconsistency errors was found in both experiments. This suggests that presentation style did not influence children's ability to correctly judge whether or not a passage made sense. Both experiments revealed developmental differences: Older children made more correct sense judgements than younger children. This is in line with the literature (e.g., Baker, 1984a, 1984b) and previous experiments in this thesis.

Also, children generally made more correct sense judgements for consistent than near and far internal inconsistency error passages. This suggests that children set appropriate standards of evaluation for error detection. However, the inclusion of an adjacent internal inconsistency passage manipulation in Experiment 7 revealed that children made a similar number of correct sense judgements for adjacent internal inconsistency, near internal inconsistency and consistent passages, and more correct sense judgements for adjacent internal inconsistency than far internal inconsistency passages. This suggests that the distance between inconsistent information influenced children's performance on the sense question. When inconsistent information is placed in adjacent sentences, with no intervening filler text, it appears that children find it easier to judge whether or not information makes sense than when information is separated by filler text. This may be because of the increased storage and processing demands associated with activating information from the evolving situation model and integrating information over several sentences. Such a finding is in line with the literature (e.g., Oakhill et al., 2005; van der Schoot et al., 2011).

Next, the reading time findings will be considered. In both experiments, reading times for the target sentence were longer when it contained an internal inconsistency error than when it was consistent. This finding indicates that children detected the internal inconsistency errors. Probably, the extra time children spent on the target error sentence reflects their effort to integrate information across sentences and compare the consistency of the information. Also, children's reading times for the target sentence were similar for the adjacent (Experiment 7 only), near and far internal inconsistency passages. This finding suggests that the distance between inconsistent information did not affect children's real-time reading behaviour. Children slowed down their processing rate to a similar extent for adjacent, near and far internal inconsistencies.

Further, in Experiment 6 there was evidence of spillover effects for older children in the analysis for all responses to the sense question. Children's reading times for the post-target sentence were longer when the passage was consistent than when it contained a near or far internal inconsistency. Thus, when analyses take into account incorrect responses, which may include instances when children were uncertain about the consistency of information, there is evidence that the processing of information from the target sentence extends into the post-target sentence. However, for the other analyses across the two experiments there was no evidence of processing spillover effects. In contrast to the analysis for all responses to the sense question, this suggests that children completed processing the error information in the target sentence before moving on to the post-target sentence. Given that previous sentences in a passage remained revealed to the reader, it may be the case that the distance between inconsistent information was not sufficiently demanding of children's information processing resources to influence the time spent reading the post-target sentence.

Finally, the relationship between monitoring performance and independent measures of word reading, reading comprehension and working memory will be considered. For the younger age group, correlations revealed that monitoring performance for adjacent internal inconsistency passages was related to reading comprehension, and for near and far internal inconsistency passages monitoring performance was related to reading comprehension and reading accuracy. However, regression analyses revealed that reading comprehension was the only significant predictor of performance for all manipulations of the internal inconsistency passages (i.e., adjacent internal inconsistency, near internal inconsistency and far internal inconsistency). This suggests that for younger children reading comprehension plays a particularly important role in the successful monitoring of internal consistency. This is perhaps unsurprising, as it is likely that reading comprehension skills are integral in achieving a full and coherent representation of the text necessary for the successful monitoring of internal inconsistency errors.

For the older age group, there were no significant moderate or strong associations between monitoring performance on adjacent internal inconsistency passages and the independent measures. Children performed well on this monitoring task (> 84%), so this finding may reflect the possibility that there was limited variance to be explained by the independent measures. However, for the near internal inconsistency passages correlations revealed that performance was related to reading comprehension, verbal working memory and numerical working memory. Whereas performance on the far internal inconsistency passages was related to numerical working memory. Regression analyses revealed that numerical working memory was the only significant predictor of performance on the sense question for both the near and far internal inconsistency passages. This suggests that for older children comprehension monitoring of internal inconsistency errors, particularly when information is separated by filler text (i.e., near and far internal inconsistency manipulations), is linked to numerical working memory capacity. Given the working memory demands associated with detecting internal inconsistency errors this is perhaps unsurprising. However, it is of interest that numerical working memory rather than verbal working memory is related to monitoring skill. This may suggest that a general working memory, rather than a modality specific linguistic working memory, plays an important role in successful monitoring of internal inconsistency errors.

To investigate the role of working memory in children's monitoring of internal inconsistency errors, the experiments included in this chapter manipulated the distance between inconsistent information. As noted in the introduction, it has been suggested that the greater the distance between inconsistent information, the greater the storage and processing demands involved in detecting the error (e.g., Albrecht & O'Brien, 1993). However, it may be the case that factors other than the distance between inconsistent information influence the working memory demands associated with detecting internal inconsistency errors, and may better capture the relationship between monitoring and working memory. Although beyond the scope of this thesis, factors that could be considered in future research are the content of intervening text and associated processing demands. Passages that include complex content with increased processing difficulty, for example temporal (order) aspects or elaboration of character goal information, are likely to impose demands on a reader's storage and processing capability because of the effort involved in integrating information into the situation model, and activating and maintaining this information during reading.

As discussed in Chapter 5, a potential limitation of the experiments included in this chapter is the relatively small number of participants included in each experiment. The issue of statistical sensitivity should be noted; low power resulting from small sample sizes increases the risk that the direction of effect and estimated effect size are inaccurate (e.g., Field, 2005).

In summary, findings revealed developmental differences in correctly judging whether or not a passage made sense. Older children made more correct sense judgements than younger children. In both experiments, children made more correct sense judgements for consistent than near and far internal inconsistency passages. Thus, suggesting that children set appropriate standards for evaluating the sense of information. Also, in Experiment 7 children made more correct sense judgements for adjacent than far internal inconsistency passages. This finding provides some support that the distance between inconsistent information influenced children's performance on the sense question. Both age groups had similar real-time reading behaviour. However, the pattern of findings was different for the two experiments, perhaps suggesting that the presentation style of passages influenced children's monitoring of internal inconsistency errors. In Experiment 6, reading times were longer for the target sentence when it contained a far internal inconsistency than when it was consistent, and reading times were similar for near internal inconsistency and consistent passages. Thus, providing some support that the distance between inconsistent information affects children's real-time reading behaviour. Whereas in Experiment 7, reading times were longer for the target sentence when it contained an internal inconsistency than when it was consistent, and reading times were similar for all internal inconsistency manipulations. Further, for young children reading comprehension and for older children numerical working memory were found to be important in the successful monitoring of internal inconsistency errors.

CHAPTER SEVEN

AN EYE-MOVEMENT EXPERIMENT INVESTIGATING COMPREHENSION MONITORING OF INTERNAL INCONSISTENCIES

The main aim of the experiment reported in this chapter was to examine children's eye-movements when reading passages containing internal inconsistencies. In addition, similar to Chapter 6, the distance between inconsistent information was manipulated to gain a greater insight into whether this aspect of the error influences monitoring behaviour.

7.1. Experiment 8

Previous experiments in this thesis have adopted the self-paced moving window method. This method has been shown to effectively tap reading comprehension processes (e.g., Just et al., 1982; Rinck, Gamez, Diaz & de Vega, 2003; van der Schoot et al., 2011). However, it should be noted that the self-paced moving window method has some limitations. One potential problem is that participants' reading rate may be slowed down because the reaction time of their eyes is faster than the reaction time of their fingers (e.g., Rayner, 1998). Another problem is that when the non-cumulative presentation method (i.e., Experiment 3: passages are presented one sentence at a time, each new sentence replaces the previous sentence) is used it prevents readers from looking back, or when the cumulative presentation method (i.e., Experiments 4 – 7: passages are presented one sentence at a time, each sentence remains on the screen) is used it prevents the experimenter from observing any look-backs. As discussed below, such limitations may be overcome by using eye-tracking methodology.

Eye-movement patterns have been used successfully to study real-time comprehension processes in reading (e.g., Rayner et al., 1989). This is because eye-movements provide experimenters with a fine-grained insight into moment-by-moment reading. One of the major advantages of eye-movement methodology over alternative measures is that it allows experimenters to distinguish between initial reading and re-reading patterns. This is important because experimenters can then determine when errors inserted into a text first influenced processing and therefore provides insight into the time course of processing error information during comprehension monitoring. Eye-movements also provide experimenters with an insight into whether or not children engage in monitoring strategies beyond reading the error sentence more slowly, such as look back and re-reading text presented earlier in the passage.

There have been relatively few studies into the nature of children's eye-movements during reading. While there has been a fair amount of research into children's comprehension monitoring of internal inconsistencies, there has been little research using eye-tracking as a methodology. Studies by Rayner (1986) and McConkie et al. (1991) provided preliminary data regarding children's basic oculomotor behaviour during reading. Both found that compared to older more proficient readers, children made longer and more frequent fixations, made more frequent regressions and had smaller perceptual spans. While the Rayner (1986) and McConkie et al., (1991) studies provide important data regarding children's eye-movement behaviour during reading, they did not investigate the eye-movement patterns of beginning and fluent readers when they process sentences containing internal inconsistencies. Indeed, very few studies to-date have set out to investigate such a question (though see van der Schoot et al., 2011).

Van der Schoot et al. (2011) investigated comprehension monitoring in 10- to 12-year-old children differing in reading comprehension skill. In Experiment 2, children's eye fixations and regressions were measured as they read narrative texts in which an action of the protagonist was consistent or inconsistent with a description of the protagonist's character given earlier. The distance between inconsistent information was also manipulated (local vs global conditions). In the local condition, both good and poor comprehenders adjusted their reading on the target sentence, having longer reading times (first-pass durations and wrap-up times) for inconsistent than consistent information. However, in the global condition poor comprehenders did not adjust their reading on the target sentence; inconsistent information did not lead to longer reading times. Findings also suggested that children were highly unlikely to go back and re-read information. Van der Schoot et al. (2011) discuss findings in the context of the situation model framework, suggesting that poor comprehenders find difficulty in constructing, not updating, a richly elaborated situation model. Given the advantages of eye-tracking, in terms of measurement sensitivity, over alternative off-line and self-paced measures, further eye-tracking experiments in the field of comprehension monitoring are needed.

To gain a greater insight into children's moment-by-moment monitoring behaviour, the current experiment recorded children's eye-movements as they read passages containing internal inconsistencies. In line with previous research (e.g., Baker, 1984a) and the findings of previous experiments, it was predicted that in comparison to older children, younger children would make fewer correct responses to the question 'did the story make sense'. It was also predicted that children would make more correct responses to the sense question for the adjacent internal inconsistency passages than for the far internal inconsistency passages. It was predicted that first pass duration would be longer for the target region when it contained an error (i.e., an internal inconsistency) than when it was consistent. This pattern would indicate real-time detection of the inconsistency during reading, and suggest comprehension monitoring has taken place. In addition, it was predicted that children would make more regressions to the target region when it contained an error than when it was consistent. This pattern would suggest that children are looking back at the source of the inconsistency, and consequently engaging in a comprehension monitoring strategy.

Method

Participants

Twenty-five children (11 boys, 14 girls) with a mean age of 7;10 years (*SD* = 3 months, range = 7;4 – 8;6 years) and 25 children (9 boys, 16 girls) with a mean age of 9;10 years (*SD* =

5 months, range = 9;4 – 10;7 years) participated. As in previous Experiments, children's word reading (measured with the TOWRE Sight Word and Phonemic Decoding) was above a standard score of 85. Children had normal or corrected-to-normal vision. Children who had special educational needs and those who spoke English as a second language were not included in the experiment. Informed consent was obtained from school headteachers and parents, and children assented to participate.

Materials

There were a total of 27 eight-sentence passages: 24 experimental passages and three practice passages. Of the 27 passages, 20 remained the same as Experiment 7 and seven underwent minor revisions (see Table 7.1 for an example and Appendix 6.1 for a list of passages). All passages were a similar length (*M* = 81.57, *SD* = 3.74, range = 75 – 88), and had a similar Flesch-Kincaid Grade readability (*M* = 2.77, *SD* = .38, range = 1.60 – 3.40). As in Experiment 7, for the experimental passages there were four versions of each passage: adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent.

Table 7.1. Example of a passage

Josh has a rabbit named Bugs.
Bugs lives in a cage in the living room, and never goes outside.
(Bugs lives in a cage in the living room, but likes going outside.)
*
Josh's Mum and Dad bought Bugs as a birthday present for Josh.
*
Bugs has grey and white fur that is very soft.
Josh feeds Bugs rabbit food and water.
*Everyday, Bugs plays in the garden on the grass.
As a special treat, Josh gives Bugs a carrot to eat every Friday.
Josh really likes having Bugs as his pet.

Note. Far internal inconsistency version presented. Adjacent internal inconsistency version indicated by the first asterisk, near internal inconsistency version indicated by the second asterisk and consistent version shown in parenthesis.

Eye-tracking method

Passages were displayed, one at a time, in black on a white background using a 17-inch monitor with a display resolution of 1024 x 768 pixels, running at 150Hz refresh rate. Monospaced 14-point font was used, and the viewing distance between each reader's eyes and the monitor was approximately 63cm. At this distance, approximately 3.2 letters of the presented text subtended 1 degree of visual angle. Viewing was binocular, but eye-movements were recorded from the right eye only using an EyeLink 1000 eye-tracking system (SR Research, Mississauga, ON, Canada) sampling at a rate of 1000Hz. Heads were positioned on a chin and forehead rest to help minimise head movements.

Procedure

Passages were assembled into four lists, each consisting of 24 experimental passages. Each list was divided into two sets (Set A and Set B) of 12 passages: three adjacent internal inconsistency, three near internal inconsistency, three far internal inconsistency and three consistent. Each list also had three practice passages: one adjacent internal inconsistency, one far internal inconsistency and one consistent passage. The same practice passages were used for each list. Passages were presented in a pseudorandomised order in which no more than two of the same type occurred consecutively within a list. Each participant saw one list, and the order of presentation for Set A and Set B was counterbalanced.

Children were tested individually at school in a quiet room, separate from their classroom. They completed the word reading (in isolation and in context) and reading comprehension measures in one 15 minute session. In two further 20 minute sessions, children completed the comprehension monitoring task (passages from Set A were presented in one session and Set B in the other session) and the working memory measures (described later). Before the start of the comprehension monitoring task, children were informed that the task was designed to examine their understanding of short stories, some of which did not make sense. Children were instructed to read silently at their normal reading pace and prepare to answer the question ‘did the story make sense’ at the end of each story.

Before the start of each session, the eye-tracker was calibrated and validated using a nine-point calibration grid. Following this, three practice passages (near internal inconsistency, far inconsistency and consistent) were presented to familiarise children with the eye-tracking equipment. Children also received feedback on their responses to the sense question to ensure they were familiar with the different passage types. After each passage, a one-point calibration was performed to correct for possible drifts in gaze position. To ensure accuracy of the gaze position, recalibration was performed following the practice passages and after every six passages. A black cross appeared after the one-point calibration on the left side of the screen above the first word of the passage. Once a stable fixation was detected within this area, the passage appeared on the screen. On each screen one passage was presented and was followed by the question ‘did the story make sense’. Children pressed the left button of a response box for ‘yes’ and the right button of the response box for ‘no’. They were not given feedback on the accuracy of their response. Once a response had been given, a new screen was presented preparing children to move on to the next story. When children were ready to continue the experimenter displayed the one point-calibration screen.

Eye fixation measures

For each passage, eye fixation measures for two regions of interest were calculated. The target region of the first piece of inconsistent information in sentence two, and the target region of the second piece of inconsistent information in sentence three, four or six depending on the error manipulation (adjacent, near and far internal inconsistency errors, respectively). Given the sensitivity of the eye-movement measure, the target region was defined as smallest amount of information required for the detection of the inconsistency error (e.g., Rayner, 1998). For the example passage shown in Table 7.1, the target region of sentence two was ‘never goes outside’ and the target region of the error sentence was ‘plays in the garden’. The target regions were defined by the experimenter and confirmed by five experienced adult readers.

The analysis that follows examines first pass duration for the target region of the error (or consistent) sentence. This is defined as the sum of the duration of all fixations on the target region from the first fixation on the target region until the first time that the reader exits the target region. This eye fixation measure reflects the initial processing of the target region. Also, the duration of readers’ first fixation after the first pass duration was measured. This was to allow possible spillover effects to be explored.

In addition, the frequency with which readers looked back to the target region of sentence two given that they had initially processed this region and moved on to at least the target region of the error (or consistent) sentence was measured (i.e., regressions). This processing measure reflects the comprehension monitoring of readers who are looking back to the source of the inconsistency, i.e., the first piece of inconsistent information provided in sentence two.

In line with the reading time experiments, for all duration measures the sum of fixations was divided by the number of syllables in the target region to yield a millisecond per syllable measure. This is because the length of the target region varied across passages.

Word reading, reading comprehension and working memory measures

Word reading (in isolation and in context) and reading comprehension measures. Children's single word and nonword reading, and reading comprehension were assessed with the TOWRE and the YARC Passage Reading, respectively (see Chapter 3 and Chapter 5 for full administration details).

Verbal and numerical working memory measures. Children's verbal and numerical working memory were assessed with the sentence span and digit span tasks, respectively (see Chapter 5 for full administration details). The sample reliabilities (Cronbach's alpha) for the verbal and numerical working memory measures were .58 and .71, respectively. As discussed previously, values in this range are comparable to the reliabilities reported for standardised measures and reliabilities reported in previous experimental work.

Results

Data filtering

The standard deviation of fixation durations was examined. No fixation durations were more than 2.5*SD* from the mean. In line with previous literature (e.g., van der Schoot et al., 2011), fixation durations less than 50ms were excluded from analysis. Also, passages that had not been read in full (i.e., children had made a response to the sense question before making at least one fixation on the last sentence of a passage) and passages that had poor calibration were excluded from analysis. This data filtering eliminated approximately 18% of the data analysed. Similar values are often reported in eye-movement studies with children (e.g., Joseph et al., 2008). Mean values for the eye-fixation measures described above were calculated over the remaining data. Before analysis the Q-Q plots for all measures were examined for each age group separately. The plots indicated that the distribution of data was acceptable.

Correct sense judgements

The mean number correct for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passage types was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The mean numbers of correct sense judgements made for each passage type by age group are shown in Table 7.2.

Table 7.2. Mean numbers (with standard deviations) of correct judgements by age group and passage type

	Age group		
Passage type	**7- to 8-year-olds (n = 25)**	**9- to 10-year-olds (n = 25)**	**Total**
Adjacent internal inconsistency (maximum = 6)	4.23 (1.32)	5.19 (1.11)	4.71 (1.22)
Near internal inconsistency (maximum = 6)	4.45 (1.37)	5.17 (1.15)	4.81 (1.26)
Far internal inconsistency (maximum = 6)	3.88 (1.82)	4.63 (1.41)	4.26 (1.62)
Consistent (maximum = 6)	5.23 (.97)	5.00 (.88)	5.12 (.93)
Total (maximum = 6)	4.45 (1.37)	5.00 (1.14)	

There was a significant main effect of age group ($F(1,48) = 6.18$, $p = .02$, $\eta p^2 = .11$) because the older children made more correct sense judgements than the younger children. There was also a main effect of passage type ($F(3,144) = 4.55$, $p = .01$, $\eta p^2 = .09$). Children made a similar number of correct sense judgements for consistent, adjacent internal inconsistency and near internal inconsistency passages (all *ts*(49) < 1.57, *ps* > .12). Children made more correct sense judgements for consistent (*t*(49) = 3.33, *p* = .01, *d* = .65), adjacent internal inconsistency (*t*(49) = 1.80, *p* = .08, *d* = .31), and near internal inconsistency (*t*(49) = 2.30, *p* = .03, *d* = .38) than far internal inconsistency passages.

There was a marginally significant interaction between age group and passage type ($F(3,144) = 2.54$, $p = .06$, $\eta p^2 = .05$), which is shown in Figure 7.1. The interaction arose because older children made a similar number of correct sense judgements for all passage types (all *ts*(24) < 1.64, *ps* > .11). Whereas younger children made more correct sense judgements for the consistent than the adjacent internal inconsistency (*t*(24) = 2.71, *p* = .01, *d* = .86), near internal inconsistency (*t*(24) = 2.19, *p* = .04, *d* = .66) and far internal inconsistency (*t*(24) = 3.57, *p* = .01, *d* = .93) passages.

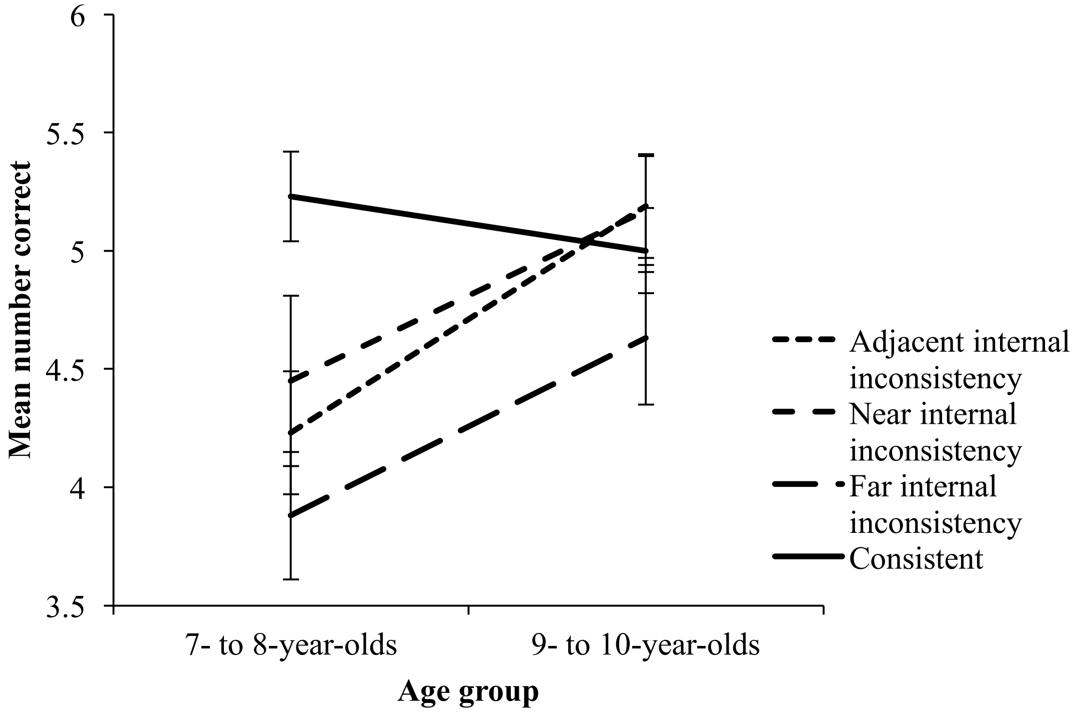

Figure 7.1. Mean numbers (+/- standard error) of correct judgements by age group and passage type

First pass duration

Similar to the reading time analyses undertaken in previous chapters, two analyses comparing the target region for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages were conducted. The first includes all responses (correct and incorrect) to the sense question; the second includes correct-only responses to the sense question.

Analysis of first pass duration for the target region for all responses (correct and incorrect) to the sense question. The mean first pass duration for the target region was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The means for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 7.3.

There was a significant main effect of age group ($F(1,48) = 7.61$, p =.01, $\eta p^2 = .14$) because the older children (M = 230.97, SD = 74.18) read the target region faster than the younger children (M = 286.03, SD = 96.12). However, there was no main effect of passage type ($F(3,144) < 1$, ns). Children had similar first pass durations for the target region when it contained an adjacent internal inconsistency (M = 259.93, SD = 90.05), near internal inconsistency (M = 259.78, SD = 75.84), far internal inconsistency (M = 263.71, SD = 77.90) and when it was consistent (M = 250.59, SD = 96.80). The interaction between age group and passage type was not significant ($F(3,144) = 1.36$, p =.26, $\eta p^2 = .03$).

Analysis of first pass duration for the target region for correct-only responses to the sense question. The mean first pass duration for the target region was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The means for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages by age group are shown in Table 7.3.

Table 7.3. Mean fixation durations (ms/syllable) (with standard deviations) for the target and spillover regions by age group, passage type and question accuracy

		Age group			
		7- to 8-year-olds		9- to 10-year-olds	
Passage type	**Question accuracy**	**Target region**	**Spillover region**	**Target region**	**Spillover region**
Adjacent internal inconsistency	**Overall**	284.57 (100.25)	255.62 (106.83)	235.28 (79.85)	215.61 (67.90)
	Correct-only	283.59 (109.91)	253.17 (139.05)	236.27 (96.35)	216.27 (70.75)[a]
Near internal inconsistency	**Overall**	281.17 (64.61)	247.26 (73.37)	238.39 (87.07)	199.80 (50.84)
	Correct-only	254.99 (63.63)	241.52 (55.00)	233.86 (90.98)	199.98 (52.76)[a]
Far internal inconsistency	**Overall**	285.59 (91.26)	216.64 (47.01)	241.83 (64.53)	212.07 (70.19)
	Correct-only	270.56 (76.65)	212.64 (70.85)	253.80 (70.74)	215.51 (78.42)[a]
Consistent	**Overall**	292.78 (128.34)	231.81 (63.46)	208.39 (65.26)	194.25 (43.86)
	Correct-only	299.92 (131.15)	234.82 (68.11)	203.18 (63.18)	190.40 (48.27)[a]

Note. Overall refers to both correct and incorrect responses. Unless indicated in the table by superscript, n = 25 for 7- to 8-year-olds and n = 25 for 9- to 10-year-olds. [a] = 24. Missing data is due to one child having no data for correct responses for the spillover region for far internal inconsistency passages.

Similar to the previous analysis there was a significant main effect of age group ($F(1,48) = 6.29$, $p = .02$, $\eta p^2 = .12$) because the older children ($M = 231.78$, $SD = 80.31$) read the target region faster than the younger children ($M = 277.27$, $SD = 95.34$). There was also no main effect of passage type ($F(3,144) < 1$, ns). In contrast to the previous analysis, the main effect of age group was qualified by a significant interaction ($F(3,144) = 3.08$, $p = .03$, $\eta p^2 = .06$), which is shown in Figure 7.2. The interaction arose because younger children had similar first pass durations for the target region for all passage types (all $ts(24) < 1.85$, $ps > .07$). Whereas older children had longer first pass durations for the near internal inconsistency ($t(24) = 2.40$, $p = .02$, $d = .39$) and far internal inconsistency ($t(24) = 3.06$, $p = .01$, $d = .75$) than the consistent passages.

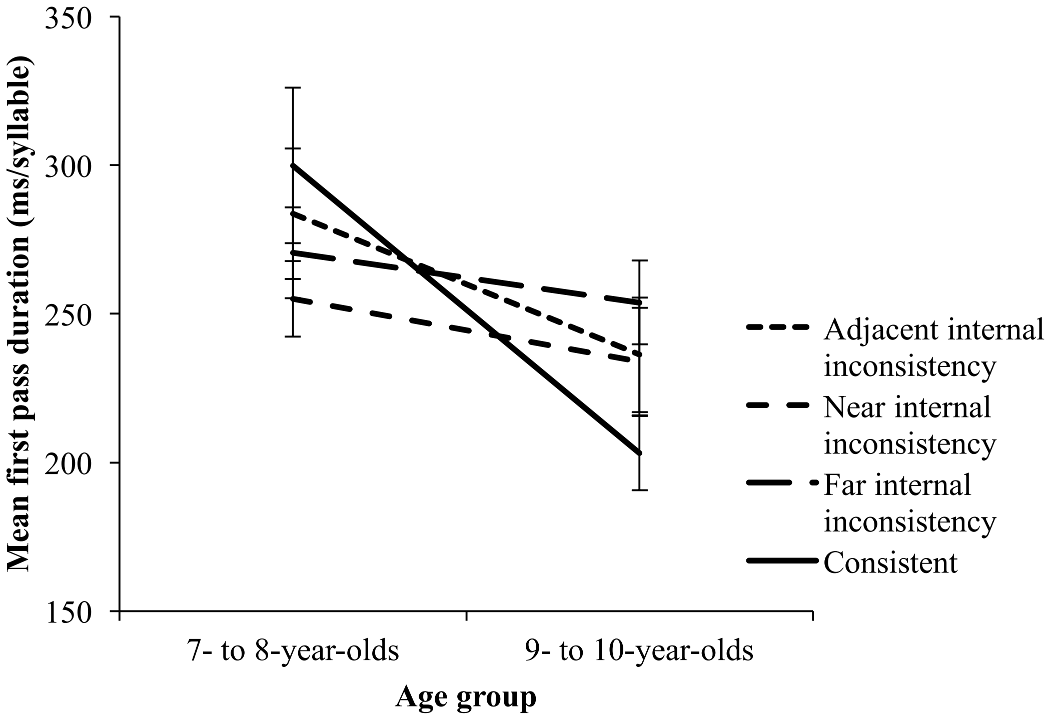

Figure 7.2. Mean first pass durations (ms/syllable) (+/- standard error) for the target region by age group and passage type

Analysis of spillover effects

Separate analyses for the spillover region for all responses (correct and incorrect) and correct-only responses to the sense question were conducted. The mean fixation duration for the spillover region was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The means for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages by age group and shown in Table 7.3.

There was a significant main effect of age group ($F(1,48) = 6.29$, $p = .02$, $\eta p^2 = .12$; $F(1,48) = 5.11$, $p = .03$, $\eta p^2 = .10$, for all and correct-only responses analyses, respectively) because the older children read the spillover region faster than the younger children. There was no significant main effect of passage type ($F(3,144) = 1.60$, $p = .19$, $\eta p^2 = .03$; $F(3,144) < 1$, ns). Further, the interaction between age group and passage type was not significant ($F(3,144) = 1.33$, $p = .27$, $\eta p^2 = .03$; $F(3,144) = 1.17$, $p = .32$, $\eta p^2 = .02$).

Regressions

Regressions to the target region of the first piece of inconsistent information (i.e., sentence two) for all responses (correct and incorrect) and correct-only responses to the sense question were analysed. The mean number of regressions for the adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent passages was the dependent variable in a 2 (age group) x 4 (passage type) ANOVA. The mean number of regressions made for each passage type by age group is shown in Table 7.4.

Table 7.4. Mean number (with standard deviations) of regressions to the target region by age group, passage type and question accuracy

		Age group	
Passage type	**Question accuracy**	**7- to 8-year-olds (n = 25)**	**9- to 10-year-olds (n = 25)**
Adjacent internal inconsistency	**Overall**	1.80 (1.32)	2.04 (1.46)
	Correct-only	1.48 (1.26)	1.76 (1.30)
Near internal inconsistency	**Overall**	1.52 (1.23)	1.40 (1.32)
	Correct-only	1.32 (1.07)	1.16 (1.11)
Far internal inconsistency	**Overall**	1.36 (1.44)	1.20 (1.35)
	Correct-Only	1.08 (1.19)	.96 (1.31)
Consistent	**Overall**	1.08 (1.47)	1.56 (1.42)
	Correct-only	.88 (1.17)	1.44 (1.36)

Note. Overall refers to both correct and incorrect responses.

The pattern of results was the same for both the analysis for all responses and correct-only responses to the sense question. There was no significant main effect of age group (both *Fs*(1,48) < 1, ns) because both age groups made a similar number of regressions. However, there was a significant main effect of passage type ($F(3,144) = 3.09$, $p = .01$, $\eta p^2 = .08$; $F(3,144) = 3.64$, $p = .01$, $\eta p^2 = .07$, for all and correct-only responses analyses, respectively). Surprisingly, children made more regressions to the target region for adjacent internal inconsistency than near internal inconsistency ($t(49) = 2.42$, $p = .02$, $d = .34$; $t(49) = 2.03$, $p = .05$, $d = .32$,), far internal inconsistency ($t(49) = 3.60$, $p = .01$, $d = .46$; $t(49) = 3.45$, $p = .01$, $d = .47$) and consistent passages ($t(49) = 2.63$, $p = .01$, $d = .42$; $t(49) = 2.30$, $p = .03$, $d = .36$). The interaction between age group and passage type was not significant ($F(3,144) = 1.05$, $p = .37$, $\eta p^2 = .02$; $F(3,144) = 1.63$, $p = .19$, $\eta p^2 = .03$).

It is clear from Table 7.4 that in both age groups children were unlikely to look back and re-read the target region of the first piece of inconsistent information. Therefore, to avoid interpretative difficulties the duration of regressions were not analysed.

Relationship between comprehension monitoring and word reading, reading comprehension and working memory

Descriptive statistics. Table 7.5 summarises the mean performance of the two age groups on independent measures of word reading, reading comprehension and working memory. Note that the performance of both age groups on the word reading and reading comprehension measures was above a standard score of 85.

Table 7.5. Mean raw scores, ability scores and standard scores (with standard deviations and t-tests) for word reading, reading comprehension and working memory measures by age group

Measure	7- to 8-year-olds (n = 25)	9- to 10-year-olds (n = 25)	$t(48)$	d
Word reading				
Raw score (maximum = 104)	62.88 (8.97)	73.24 (6.53)	4.67^{***}	1.32
Standardised score	116.16 (9.37)	111.96 (9.29)	1.59	.45
Nonword reading				
Raw score (maximum = 63)	33.92 (10.61)	42.60 (6.97)	3.42^{**}	.96
Standardised score	115.40 (11.80)	115.08 (10.56)	< 1	-
Reading rate				
Raw score (seconds)	99.46 (25.37)	93.28 (16.24)	1.03	.29
Ability score	66.56 (9.57)	81.28 (7.94)	5.92^{***}	1.67
Standardised score	111.76 (9.02)	113.84 (10.69)	< 1	-
Word reading accuracy				
Raw score	4.86 (2.15)	2.78 (1.58)	3.89^{***}	1.10
Ability score	54.08 (14.47)	64.12 (9.07)	2.94^{**}	.83
Standardised score	108.36 (9.92)	111.04 (12.14)	< 1	-
Reading comprehension				
Raw score (maximum = 8)	4.94 (1.32)	9.46 (1.14)	1.49^{*}	.42
Ability score	56.20 (7.14)	64.28 (8.90)	3.54^{**}	1.00
Standardised score	107.76 (8.67)	107.72 (11.72)	< 1	-
Verbal working memory				
Number of items (maximum = 42)	12.72 (5.25)	21.24 (6.69)	5.01^{***}	1.42
Number of trials (maximum = 12)	3.80 (1.22)	5.64 (1.44)	4.86^{***}	1.38
Span (maximum = 5)	2.28 (.54)	3.04 (.61)	4.65^{***}	1.32
Numerical working memory				
Number of items (maximum = 27)	14.08 (5.75)	18.52 (5.53)	2.78^{**}	.79
Number of trials (maximum = 9)	4.28 (1.65)	5.52 (1.78)	2.56^{*}	.72
Span (maximum = 4)	2.40 (.71)	2.88 (.73)	2.37^{*}	.67

Note. Raw scores are presented on the upper row, and where appropriate ability scores on the middle row and standardised scores on the lower row. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Interrelations between measures. The same correlation analyses were conducted as in the previous chapters. Here the correlation between the TOWRE Sight Word reading and TOWRE Phonemic Decoding measures was *rs* = .80 and .54, *ps* < .01, for older and younger age groups, respectively. Although the correlation for the younger age group is a little low, a composite word and nonword reading measure was created so that the analysis was in line with previous chapters. Also, the pattern of results was similar for both the composite and individual measures. Note, despite the moderate association between verbal and numerical working memory measures in both age groups (see Table 7.6 and 7.7 for values), a composite measure was not created so that the independent relations for each measure could be explored. Measures of monitoring performance were accuracy to the sense question for the adjacent, near and far internal inconsistency passages. Consistent passages were not included in analysis because performance was close to ceiling for both age groups.

For brevity the reader is directed to Tables 7.6 and 7.7 which summarise the correlations for each age group, and only key findings are discussed here. As noted in Chapter 5, many correlations did not reach statistical significance and so will be interpreted in relation to effect sizes. For the 7- to 8-year-old children, there was a moderate association between correct responses to the sense question for the adjacent internal inconsistencies and the word and nonword reading measure. For the younger age group, there was also a moderate association between correct responses to the sense question for the far internal inconsistencies and the verbal working memory measure. However, for the 9- to 10-year-old children, there were no moderate or strong associations between correct responses to the sense question for adjacent, near or far internal inconsistency passages and the other measures.

Table 7.6. Correlations between measures by age group: 7- to 8-year-olds (n = 25)

	1	2	3	4	5	6	7	8
1. Word and nonword reading	-	.37	-.12	.24	.22	.47*	.04	.33
2. Word reading accuracy	.37	-	-.36	-.04	.04	.33	-.38	-.07
3. Reading comprehension	-.12	-.36	-	.13	-.44*	.07	.12	.28
4. Verbal working memory	.24	-.04	.13	-	.38	.09	.39	.42*
5. Numerical working memory	.22	.04	-.44*	.38	-	.08	.25	-.02
6. Sense question: Adjacent internal inconsistency	.47*	.33	.07	.09	.08	-	.39	.31
7. Sense question: Near internal inconsistency	.04	-.38	.12	.39	.25	.39	-	.42*
8. Sense question: Far internal inconsistency	.33	-.07	.28	.42*	-.02	.31	.42*	-

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 7.7. Correlations between measures by age group: 9- to 10-year-olds (n = 25)

	1	2	3	4	5	6	7	8
1. Word and nonword reading	-	.66***	.32	.10	-.10	.37	-.05	-.15
2. Word reading accuracy	.66***	-	.76***	.20	.22	.29	.07	-.03
3. Reading comprehension	.32	.76***	-	.07	.06	.23	.24	.35
4. Verbal working memory	.10	.20	.07	-	.30	-.26	-.27	-.15
5. Numerical working memory	-.10	.22	.06	.30	-	-.13	-.06	-.30
6. Sense question: Adjacent internal inconsistency	.37	.29	.23	-.26	-.13	-	.52**	.10
7. Sense question: Near internal inconsistency	-.05	.07	.24	-.27	-.06	.52**	-	.17
8. Sense question: Far internal inconsistency	-.15	-.03	.35	-.15	-.30	.10	.17	-

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

The role of word reading, reading comprehension and working memory in monitoring ability. As in the previous chapters, forward stepwise regressions were conducted to explore the specific role of word reading, reading comprehension and working memory in comprehension monitoring ability. The dependent variables were performance on the sense question for adjacent and far internal inconsistency passages. Because there were no significant moderate or strong associations between the monitoring measures and the other measures in the older age group, regressions were only conducted for the younger age group. In the younger age group, there were no significant moderate or strong associations between the near internal inconsistency passages and the other measures, so regressions were only conducted for the adjacent and far internal inconsistency passages. As before, the predictors entered were: word and nonword reading, reading comprehension, verbal working memory and numerical working memory.

For adjacent internal inconsistencies, the model explained 25% of variance in performance on the sense question ($F(1,23) = 7.55$, $p = .01$). Word and nonword reading was the only significant predictor of performance ($\beta = .50$, $p = .01$). Whereas for far internal inconsistencies, the model explained 19% of variance in performance on the sense question ($F(1,23) = 5.48$, $p = .03$). Verbal working memory was the only significant predictor of performance ($\beta = .44$, $p = .03$).

Discussion

The main aim of Experiment 8 was to examine 7- to 10-year-old children's eye-movements when reading passages containing internal inconsistencies in which the distance between inconsistent information was manipulated. First, findings for the sense question will be considered. As predicted, and in line with previous experiments in this thesis, developmental differences were found: Older children made significantly more correct sense judgements than younger children.

In general, children made more correct sense judgements for consistent, adjacent internal inconsistency and near internal inconsistency than far internal inconsistency passages. This suggests when inconsistent information is separated by filler text, children struggle to accurately judge whether or not information makes sense. This may be because of the increased processing and storage demands associated with activating information from the evolving situation model and integrating information over several sentences.

However, the distance between inconsistent information does not seem to influence performance on the sense question when the influence of age group is taken into consideration. The marginally significant interaction between age group and passage type revealed that the two age groups differ in their performance on the different passage types. Older children made a similar number of correct sense judgements for all passages types. Whereas, younger children made more correct sense judgements for the consistent than the adjacent internal inconsistency, near internal inconsistency and far internal inconsistency passages. This suggests that even when information was separated by filler text, older children were able to set appropriate standards of evaluation for detecting internal inconsistency errors. However, younger children experienced difficulties in correctly judging that inconsistent information did not make sense irrespective of the distance between inconsistent information. Thus, it seems that young children struggle to adopt the internal consistency standard.

Next, the eye-movement findings will be considered. For the correct-only analysis there was a significant interaction between age group and passage type. This revealed that younger children spent a similar amount of time reading the target region of the different passage types (adjacent internal inconsistency, near internal inconsistency, far internal inconsistency and consistent). The general increase in reading times for the target region displayed by younger children may suggest that these children need to undertake a more purposeful and careful reading of passages that may contain internal inconsistency errors in order to correctly detect whether or not they make sense. As mentioned above for the sense questions, this may reflect young children's difficulty in monitoring their sense for internal consistency. In contrast, older children adjusted their reading times for the target region depending on the distance between inconsistent information. Older children only spent more time reading the target region when the distance between inconsistent information was increased in the near and far internal inconsistency passages. This finding may reflect the additional processing demands involved in integrating information across several sentences and evaluating the sense of information.

In addition, there was no evidence of spillover effects. This suggests that children completed processing information in the target region before moving on and reading additional text. For younger children, such a suggestion is in line with the idea that they are undertaking a purposeful and careful reading of the text, which includes integrating information and evaluating the sense of information before moving on to new information. Whereas, for older children it may suggest that the integration processes involved in detecting inconsistent information were not sufficiently demanding of information processing resources to influence reading times for regions beyond the target region.

Surprisingly, the regression analysis revealed that children in both age groups were unlikely to go back and re-read the target region of the first piece of inconsistent information. This finding is in contrast to predictions. However, previous eye-movement research has found both the presence and absence of regressions in response to internally inconsistent information (e.g., Poynor & Morris, 2003; van der Schoot et al., 2011). It is possible that the type of contradictory information used to create the internal inconsistency errors did not compel children to reinspect the first piece of inconsistent information. Typically, inconsistent information related to character or object information (e.g., a rabbit that never goes outside, but plays in the garden). Once this type of information is integrated into the situation model, children may not experience a feeling of misunderstanding that compels them to question the validity of this information and look back to earlier parts of the text. Such an argument may also explain the finding that children made more regressions for adjacent internal inconsistency passages. When inconsistent information is presented in adjacent sentences it may be the case that children do not have sufficient time for the information to be integrated and established within their situation model. Thus, when children encounter contradictory information they may be unsure whether or not they misread or misunderstood the earlier information, so look back to check.

Finally, the relationship between monitoring performance and independent measures of word reading, reading comprehension and working memory will be considered. For the younger age group, correlations and regression analyses revealed that monitoring performance for adjacent internal inconsistency passages was related to word and nonword reading, and for far internal inconsistency passages monitoring performance was related to verbal working memory. This suggests that correctly judging the sense of inconsistent information is more closely associated with different skills depending on the distance between inconsistent information. When information is in adjacent sentences word and nonword reading skills are important, however, when information is separated by filler text verbal working memory capacity is important. Successful monitoring of internal inconsistencies can be demanding for young children. Thus, greater word reading proficiency allows children to allocate more processing resources to monitoring processes (e.g., Perfetti, 1985), and increased working memory capacity facilitates the storage and processing of information necessary for the detection of inconsistent information that is separated by filler text (e.g., Oakhill et al., 2005).

For the older age group there were no moderate or strong associations between performance on the monitoring measures and the independent measures. As noted in previous chapters, children performed well on the monitoring task (> 77% for all passage types). So this finding may reflect the possibility that there was limited variance to be explained by the independent measures.

A strength of the experiment included in this chapter is the use of eye-tracking methodology. In comparison to the off-line (Experiment 1: pen and paper task) and other real-time (Experiments 3 – 7: self-paced moving window) methods used in this thesis, eye-tracking can be considered a more sensitive measure of reading. As noted in the introduction, one of the major advantages of eye-movement methodology over alternative measures is that it allows experimenters to distinguish between initial reading and re-reading patterns. This is important because experimenters can then determine when errors inserted into a text first influenced processing and therefore provides insight into the time course of processing error information during comprehension monitoring. Eye-movements also provide experimenters with an insight into whether or not children engage in monitoring strategies beyond reading the error sentence more slowly, such as look back and re-reading text presented earlier in the passage.

In summary, findings revealed developmental differences in correctly judging whether or not a passage made sense. Generally, older children made more correct sense judgements than younger children. Older children made a similar number of correct sense judgements for all passage types, whereas younger children made more correct sense judgements for consistent than internal inconsistency passages. Thus, the distance between inconsistent information did not seem to influence performance on the sense question. In contrast, first pass durations revealed that the older children adjusted their reading times for the target region depending on the distance between inconsistent information. Older children spent more time reading the target region of near and far internal inconsistency passages. These findings suggest that older children have developed the skills necessary to successfully monitor their comprehension for internal consistency, whereas these skills are still developing in younger children. In both age groups, there were few instances of children looking back and re-reading the target region of the first piece of inconsistent information. Further, for young children in particular, word and nonword reading and verbal working memory were found to be important in the successful monitoring of internal inconsistency errors.

CHAPTER EIGHT

DISCUSSION AND CONCLUSIONS

The aim of the research presented in this thesis was to explore the development of comprehension monitoring in children aged 7 to 10 years, and the critical task and reader characteristics that influence this skill. This discussion begins with a summary of the experimental findings in relation to this aim. Limitations of the experimental work and implications for our understanding of comprehension monitoring, education and future research are then presented.

8.1. Comprehension monitoring: Developmental differences

By comparing the performance of 7- to 10-year-old children on a range of comprehension monitoring tasks, developmental differences in comprehension monitoring were explored. The following is a summary of the experimental findings related to developmental differences in comprehension monitoring. First differences associated with task characteristics will be presented, followed by reader characteristics.

8.1.1. Task characteristics

This section will discuss findings related to two task characteristics: the type of error inserted into passages and the extent to which instructions alert participants that texts contain errors. Each will be discussed in turn.

Error Type

The type of error inserted into texts can influence comprehension monitoring (e.g., Baker, 1984a). This is because the detection of different errors requires readers to adopt different standards of evaluation, and each standard is associated with different processing demands. One reason for developmental differences in comprehension monitoring is that children's knowledge of different standards, and their ability to use them proficiently may develop throughout childhood. In this work, children's monitoring of three error types was explored: nonwords, general knowledge violations and internal inconsistencies.

Nonword

Performance in both age groups was near to ceiling for nonword errors. Children were able to correctly judge that information containing nonwords did not make sense and accurately identify this information in text. Children also made more correct sense judgements and error identifications for nonword than general knowledge violations and internal inconsistencies. These findings suggest that 7- to 10-year-old children are able to proficiently adopt the lexical standard, and this standard is easier to adopt than the external and internal consistency standards. This is perhaps unsurprising given the low processing demands associated with checking that a word is in one's mental lexicon versus comparing information from the text world knowledge or evaluating the consistency of just reading information with previously read information.

General knowledge violation

Developmental differences in children's monitoring of general knowledge violations were found. Older children were better than younger children at correctly judging that information violating world knowledge did not make sense and accurately identifying this information in text. The error identification information revealed that younger children's poorer performance was not due to an inability to undertake monitoring behaviours and identify error information. Rather, it seems that younger children undertake similar monitoring behaviours to older children, but they are less accurate.

In general, children made more correct sense judgements for general knowledge violations than internal inconsistencies. This suggests children find it easier to monitor the sense of information when it involves prior knowledge external from the text. It is likely that a representation of this information is more stable in the reader's mind, so contradictory information may be more apparent and less likely to be accepted as accurate.

In addition, the explicitness of error information was found to influence children's ability to correctly judge whether or not information made sense. In both age groups, children were more likely to correctly judge that information that was unrelated to the standard state of affairs did not make sense, than information that was closely related to the standard state of affairs. However, children's reading times were similar for related and unrelated information. This suggests that the explicitness of error information did not influence children's real-time reading behaviour. Perhaps comparing information from the text with world knowledge imposes similar processing demands irrespective of how closely related to the standard state of affairs that information is. Also, the pattern of reading time findings was the same for the analyses that included all responses (and thus regardless of accuracy) and correct-only responses to the sense question. This suggests that children generate an internal signal of error detection during reading, but sometimes fail to act on this signal during post-reading activities (e.g., responding to a sense question). Furthermore, there were no spillover effects for general knowledge violation passages. This suggests that children's processing of external consistency errors does not extend beyond the error sentence. Thus, it may be the case that when monitoring their comprehension for the types of general knowledge violations included in this work, 7- to 10-year-old children did not encounter sufficient demands in terms of monitoring skills or processing resources to spend additional time considering error information.

Internal inconsistency

Development differences were found in children's monitoring of internal inconsistency errors. Older children were better than younger children at correctly judging that contradictory information did not make sense and accurately identifying this information in text. In comparison to nonwords and general knowledge violations, children in both age groups encountered most difficulties in monitoring their comprehension for internal consistencies. It is likely that this is because detection of internal inconsistency errors requires readers to integrate information across sentences and construct a rich and coherent situation model. These reading behaviours can be considered demanding of children's monitoring skills and information processing resources.

In this work, the processing demands associated with monitoring internal inconsistencies were manipulated by increasing the distance between inconsistent information. Inconsistent information was presented in adjacent sentences or separated by one or three filler sentences (adjacent, near and far internal inconsistency versions, respectively). The distance between inconsistent information did not seem to influence performance on the sense question. Older children made a similar number of correct sense judgements for all passage types (internal inconsistency and consistent). Whereas younger children made more correct sense judgements for the consistent than the inconsistent passages, but a similar number of correct sense judgements for the inconsistent passages. This suggests that irrespective of the distance between inconsistent information older children are able to set appropriate standards for evaluating its consistency. However, proficiency in this skill is still developing in younger children.

In contrast to the sense question data, the reading time data revealed that children in both age groups had similar patterns of reading behaviour. In general, children spent more time reading internally inconsistent than consistent information, and a similar amount of time reading the inconsistent information irrespective of the distance between contradictory information. This suggests that the distance between inconsistent information did not affect children's reading behaviour, and children experienced similar processing demands when reading the different passage types. However, when a more fine-grained measure of processing time was assessed (first pass duration) in the eye-movement experiment developmental differences in reading behaviour were found. Younger children spent a similar amount of time reading the target region irrespective of passage type. Whereas older children only spent more time reading the near and far internal inconsistency passages. This finding may suggest that 7- to 10-year-olds monitor their comprehension for internal consistency in different ways. Younger children engage in a more thorough and purposeful reading of the entire text, whereas older children adjust their reading depending on the distance between inconsistent information. This increased reading time may reflect the additional processing demands involved in integrating information across several sentences.

The eye-movement experiment also revealed that children in both age groups were unlikely to look-back and re-read the first piece of inconsistent information. This may be due to the type of inconsistent information included in passages or the limited distance between inconsistent information (see Section 8.2 below for a further discussion of these limitations).

There was some evidence of spillover effects (Experiment 6), which may suggest that older children spend longer reading the post-target sentence when they are uncertain about the consistency of information. In the main, the data did not provide any evidence for spillover effects for internal inconsistency passages. As mentioned above, this suggests that children's processing of internal consistency errors does not extend beyond the error sentence. Because children were provided with instructions that alerted them to the presence of errors in texts and their task of judging the sense of passages, they may have adjusted their reading behaviour. Children may have undertaken a more purposeful and careful reading of passage, which included evaluating the sense of each sentence before moving on and reading new information.

In summary, developmental differences in comprehension monitoring were found. Older children made more correct sense judgements and accurate error identifications than younger children. However, children in both age groups displayed a similar pattern of reading time findings. Together this suggests a quantitative rather than a qualitative change in the monitoring skills of beginning and proficient readers. It seems that young children have similar monitoring skills to older children, but these skills become more proficient over time and children become more adept at interpreting the meaning of internal detection signals generated during reading. Furthermore, the type of error influenced monitoring success. Children were better at monitoring their comprehension for nonword than general knowledge violation and internal inconsistency errors. The explicitness of error information and to some extent the distance between inconsistent information also influenced monitoring performance.

Instructions

Another task characteristic relevant to comprehension monitoring is the extent to which instructions given to participants alert them to the presence of errors in texts. Developmental differences in comprehension monitoring may be related to differences in children's perceptions of reading, and consequently the goals they set when reading (e.g., Cross & Paris, 1988; Myers & Paris, 1978). Younger children may attend less to the meaning processes in reading than older children, and therefore fail to evaluate their understanding of information. Instructions that inform children about their task of judging the sense of information may influence their processing of text and engagement in monitoring behaviours. In Experiment 3, children completed two conditions with different instructions: they were either alerted or not alerted to the presence of errors in texts. The same pattern of findings was found for both age groups. Even when children were not alerted to the presence of errors in texts, reading times for the target error sentences were generally longer than consistent sentences. This suggests that children in both age groups attended to the meaning of information and monitored its sense. In addition, children had longer reading times for the target sentences when they were alerted that texts may contain errors, and similar reading ties for consistent and inconsistent text. This suggests that instructions that alert children to the likely presence of errors in texts change reading behaviour and this is particularly true for more complex errors. One reason for this is that children undertake a more purposeful and careful reading of the text, which fosters higher levels of evaluation and engaging in monitoring strategies (e.g., double-checking the sense of information).

8.1.2. Reader characteristics

This section will focus on the reader characteristics of word reading, reading comprehension and working memory. This work explored the relationship between comprehension monitoring and word reading, reading comprehension and working memory by including independent measures of each skill. First, general comments relating to findings for the independent measures will be presented. Second, findings for each characteristic will be presented.

There was a different pattern of findings for general knowledge violation and internal inconsistency errors. This is perhaps unsurprising given the different skills involved in evaluating external consistency compared to internal consistency. There was also a different pattern of findings for older and younger children. One reason for this may be that monitoring success is related to different skills during the course of development. Another reason may be that the older children showed high levels of performance on the monitoring tasks, so there is limited variance to be explained by the independent measures. Specific findings related to word reading, reading comprehension and working memory are discussed below.

Word reading is related to comprehension level in the early years (e.g., Juel et al., 1986). For this reason, it is likely that young children's ability to proficiently monitor their comprehension may be influenced by their word reading skill. In this work, skill in monitoring general knowledge violations was related to word and nonword reading in the younger children. As suggested by Perfetti (1985) proficient word reading frees up processing resources. Given that younger children are more likely to have lower processing capabilities; it may be the case that they benefit most from the availability of additional processing resources, which can be allocated to higher levels of processing such as comprehension monitoring.

Comprehension monitoring is an important component process of reading comprehension, helping to ensure that a consistent and coherent representation of a text is constructed (e.g., Cain, Oakhill & Bryant, 2004; Rubman & Waters, 2000). This work did not find a strong relationship between comprehension monitoring and reading comprehension. Across experiments findings were mixed: no relationship was found between skill in monitoring general knowledge violations and reading comprehension, whereas a relationship was found between skill in monitoring internal inconsistencies and reading comprehension. It may be the case that reading comprehension is more closely associated with monitoring internal consistency because the processes involved in these two behaviours are similar. Both processes require readers to construct a representation of the text, including integrating and updating information. Whereas, monitoring external consistency only requires readers to check that information is consistent with general world knowledge. Nonetheless, given the role that comprehension monitoring is thought to play in constructing a consistent and coherent representation of the text, finding so little relationship between comprehension monitoring and reading comprehension is perhaps surprising. However, this finding may be accounted for by limitations in statistical sensitivity that result from the relatively small sample sizes included in the independent measures analyses (see the critical evaluation of the experimental work, Section 7.2, for a full discussion of this issue).

Working memory capacity is related to comprehension monitoring ability (e.g., Oakhill et al., 2005; van der Schoot et al., 2011). It is important because it allows readers to establish co-reference, integrate and update information, so that an adequate representation of the text can be constructed. For this reason, it is thought that working memory may be particularly important for successfully monitoring internal consistency. Also, working memory capacity increases throughout childhood (e.g., Gathercole, 1998). Thus, older children may have greater storage and processing capacities, which facilitate monitoring. In this work, skill in monitoring general knowledge violations was not related to working memory. Given that there are limited storage and processing demands associated with readers checking the correspondence between information presented in a text and their general world knowledge, working memory may not be expected to strongly relate to successful monitoring of general knowledge violations.

In contrast, for the internal inconsistencies there was some evidence that monitoring skill was related to working memory. Findings revealed that both verbal and numerical working memory was associated with skill in monitoring internal inconsistencies, particularly inconsistencies where information was separated by intervening filler text. However, no clear pattern of findings was found. Interestingly, it seems that although working memory capacity may be critically involved in many language processes (e.g., Daneman & Merikle, 1996), there is only a weak relationship between comprehension monitoring and working memory capacity. Thus, suggesting that the monitoring process does not rely on a reader's ability to store and process information. This finding questions the importance of working memory capacity for monitoring proficiency, and working memory capacity as a source of monitoring difficulties. However, as mentioned above for reading comprehension, it may be the case this finding can be accounted for by low statistical sensitivity resulting from relatively small sample sizes. Also, it is possible that methodological limitations such as the distance between inconsistent information, the complexity of passage content, and the inclusion of a verbal component in the numerical working memory measure meant that measures did not adequately tap working memory (see the critical evaluation of the experimental work, Section 8.2, for a full discussion of these issues).

In summary, this work assessed the relationship between comprehension monitoring and word reading, reading comprehension and working memory. No clear relationship between comprehension monitoring and the independent measures was found. This may suggest that the skills measured are only weakly related to the ability to monitor comprehension, and skills beyond those assessed here are more closely associated with monitoring skill. However, methodological limitations, such as relatively small sample sizes, low statistical power, and the appropriateness of materials and measures (discussed below in Section 7.2) may also account for for findings.

8.2. Ways to assess comprehension monitoring

This section will discuss the findings related to the use of off-line and real-time measures. Experiments in this thesis included off-line measures, specifically underlining of error information and responses to the question 'did the story make sense?', and real-time measures, specifically self-paced reading times and eye-movements. Both off-line and real-time measures demonstrate sensitivity to comprehension monitoring behaviours.

Verbal responses to the sense question provided valuable information about whether or not children had evaluated the sense of text. The underlining of text, reading times and eye-fixations provided additional information about whether or not children had detected the source of miscomprehension. The underlining of error information provided a measure of error identification, however the act of underlining problematic information during reading lacks ecological validity. Self-paced reading times provided an insight into the comprehension monitoring processes that occur during reading, and whether or not encountering erroneous information influenced reading behaviour. Eye-tracking methodology offered clear advantages over both off-line and self-paced reading time measures. Eye-movements provided high temporal resolution allowing the distinction between initial reading and re-reading patterns. This meant that it was possible to determine when errors inserted into a text first influenced processing, and therefore gain an insight into the time course of processing error information during comprehension monitoring. Eye-movements also provided an insight into whether or not children engaged in monitoring strategies beyond reading the error sentence more slowly, such as looking back and re-reading text presented earlier in the passage.

In summary, off-line and real-time measures provide different information about the monitoring process; behaviours which occur after and during the reading process, and responses demand different skills from participants. Comprehension monitoring is a multi-dimensional process, thus to gain a true insight into monitoring behaviour there is a need for research to include both reading time measures (self-paced or eye-movement) and verbal responses to comprehension questions.

8.3. A critical evaluation of the experimental work

There were limitations to the experimental work. This section focuses on restrictions based on the sample, weaknesses of the error information, potential bias created by task instructions and suitability of the independent measures. These limitations are discussed in turn.

The first limitation relates to the relatively small number of participants included in the experiments. Sample sizes similar to those included in the experiments in this thesis are typical in developmental studies investigating the psychology of language (e.g., Oakhill et al., 2005; van der Schoot et al., 2011). Nonetheless, small sample sizes reduce statistical power and increase the risk of inaccurately interpreting the direction and size of any effects (e.g., Field, 2005). The issue of low statistical power is particularly noteworthy in the correlation analyses included in this thesis. Given the relatively small size of samples included in the correlation analyses, findings are difficult to interpret. Although beyond the scope of this thesis, further analysis of correlations, which combines data on subject scores across experiments, could be undertaken. Also, future research should consider the utility of larger samples.

Another limitation related to the sample is the restricted age range of the participants included in this work. All experiments compare the performance of two age groups (7- to 8-year-olds and 9- to 10-year-olds) that were considered to be representative of the transition between beginning and fluent reading. However, it may be the case that critical stages in the development of monitoring skills occur outside of this age range. Such differences will not be captured by the work included in this thesis. Longitudinal work would allow the developmental trajectory of monitoring skills to be established and related to specific skills. However, such work was beyond the scope of this thesis.

A further limitation relates to the nature of the error information inserted into the related and unrelated general knowledge violation passages and internal inconsistency passages. First, the general knowledge passages will be considered (Experiments 4 and 5). The explicitness of error information was decided by the experimenter and verified by a sample of experienced adult readers, but the opinions of children were not measured. It may be the case that children's opinion on the explicitness of information differs from that of adults. Differences between correct responses to the sense question for related and unrelated general knowledge violations were found. So, it is likely that children perceived some differences in the explicitness of information. However, it may be the case that they did not consider the information to differ to the extent perceived by the experimenter. To overcome this potential limitation, during the pilot phase of future work it may be advantageous to measure children's opinion on the explicitness of information. This could be achieved by providing children with a number of words that are both related and unrelated to the standard state of affairs, and asking children to order words from the most to the least related.

Second, the internal inconsistency errors will be considered. This work focuses on logical internal inconsistency errors in which two pieces of information contradicted one another. Typically, the contradictory information involved character or object descriptions. In comparison to other types of contradictory information that concerns temporal (order) aspects or requires the elaboration of character goal information, it may be easier to interpret character or object descriptions, integrate this information into the situation model, and activate and maintain this information during reading. However, this work does not manipulate the type of contradictory information inserted internal inconsistency errors. So no definite conclusions can be drawn. To address this, future research could explore the development of children's monitoring of different types of contradictory information, and the specific skills related to monitoring of different contradictory information.

In addition, the distance between inconsistent information should be considered. For Experiments 6-8 it may be the case that the distance between contradictory information was not sufficient to ensure that the first piece of inconsistent information was no longer active in working memory. Previous research has demonstrated effects of distance when inconsistent information is separated by two or three sentences (e.g., Ackerman, 1984a, 1984b; Glanzer, Dorfman & Kaplan, 1981; Oakhill et al., 2005), and this work also found some effects of distance. However, such effects may be magnified if the distance between inconsistent information was increased. Alternatively, it may be the case that factors other than the distance between contradictory information influence the working memory demands associated with monitoring internal consistency. Factors such as the type of contradictory information (as discussed above) and the content of intervening text may better capture the relationship between monitoring and working memory. Such information is likely to impose demands on a reader's storage and processing capability because of the effort involved in integrating information into the situation model, and activating and maintaining this information during reading. To address this, future research could manipulate the content of intervening text, and its associated processing demands, as a means of exploring the effective component in the relationship between monitoring internal consistency and working memory.

A further limitation of this work relates to the instructions provided to children when undertaking the monitoring tasks. With the exception of Experiment 3, experiments in this thesis provided children with instructions that alerted them to the presence of errors in texts. Research has suggested that such instructions may be necessary to gain an accurate insight into comprehension monitoring (e.g., Baker, 1985; Zabrucky & Moore, 1989). However, as demonstrated by Experiment 3, task instructions can influence monitoring performance and reading behaviours. Thus, findings from this work should be considered in light of this instruction manipulation.

There may also be a limitation related to the independent measures included in this work. The data did not reveal a consistent pattern of findings for the relationship between monitoring skill and the independent measures. As noted previously, this may be due to small sample sizes, low statistical sensitivity or limited variance to be explained. However, it is possible that alternative measures may have more appropriately tapped non-linguistic working memory. Digit span tasks were used to tap non-linguistic working memory. These tasks have been used previously as a measure of non-linguistic working memory in comprehension monitoring research with large samples of school age children (e.g., Oakhill et al., 2005; Seigneuric et al., 2000). However, these tasks require children to read and recall digits. Thus, there is a verbal component to the tasks. A non-linguistic working memory measure that does not involve a verbal component, such as the spatial memory span test (e.g., Nation, Adams, Bowyer-Crane & Snowling, 1999) may more appropriately tap non-linguistic working memory.

Further, it is possible that the independent measures included in this work did not tap all of the skills important for monitoring success. Given the restrictions on the amount of time that it is reasonable to work with one child in a school environment, only a select number of what were theoretically considered to be measures of important independent skills were chosen to accompany the monitoring tasks. This work did not include a measure of attentional control. Research has suggested that attentional control may play an important role in successful comprehension monitoring (e.g., Arrington, Kulesz, Francis, Fletcher & Barnes, 2014; Brown, 1987). This is because attentional control is associated with the ability to inhibit or suppress interfering, irrelevant or prepotent responses, and to initiate those processes that are more relevant while maintaining attention on task-relevant information (e.g., Conners, 2009; Gathercole & Baddeley, 1993). It is likely that such processes will impact the integration and updating of information – critical aspects of comprehension monitoring – and promote the development of a coherent representation of a text. However, to-date the relationship between comprehension monitoring and attentional control remains unaddressed. In future research it may be insightful to include a measure of attentional control to explore the relation between this skill and comprehension monitoring.

8.4. Implications

The section that follows provides an evaluation of the findings in relation to the contributions to comprehension monitoring research, implications for education and implications for future research. First, contributions to comprehension monitoring research are considered. Second, implications for education are considered. Third, implications for future experimental work are considered.

8.4.1. Contributions to comprehension monitoring research

The work included in this thesis provides valuable contributions to our understanding of comprehension monitoring. This section summarises the contributions in relation to our understanding of the development of monitoring, the nature of the monitoring process, reading strategies and standards of evaluation adopted when monitoring, and the relationship between monitoring and working memory capacity.

Many previous investigations into comprehension monitoring have focused their attention on the effects of reading and comprehension ability on children's evaluation of comprehension. This work explores comprehension monitoring developmentally during the period of transition between beginning and fluent reading. Thus, novel insight into the development of comprehension monitoring is provided. Findings established developmental differences in comprehension monitoring of 7- to 10-year-old children. Older children were better at correctly judging the sense of information and more likely to adjust their reading behaviour in relation to error information. However, both age groups displayed similar reading behaviour. This suggests that young children undertake similar monitoring behaviours, albeit with different levels of success.

The nature of the monitoring process remains a controversial topic in monitoring research. The work on task instructions adds to our understanding of the nature of the monitoring process. Children monitored their comprehension without being instructed to do so. This finding suggests that children spontaneously adopt an 'evaluative mindset' and that comprehension monitoring is a passive process that children routinely engage in when reading. This work, in conjunction with findings already appearing in the literature (see Singer, 2013 for a review) contributes evidence to the debate whether monitoring is a passive or strategic process in favour of the former. Whilst comprehension monitoring may be a passive process, this work suggests that monitoring can be influenced by task and reader characteristics. One important task characteristic is the extent to which task instructions alert readers to the likely presence of errors in texts. This work revealed that when provided with instructions that alert to the presence of errors in texts, children undertake a more purposeful and careful reading of the text which fosters higher levels of evaluation and the adoption of critical monitoring strategies.

This work also extends our knowledge of the reading goals, standards of evaluation and monitoring strategies that children adopt when monitoring their comprehension. Findings revealed that children were able to adopt lexical, external consistency and internal consistency standards. Adopting such standards requires children to integrate and compare the comprehensibility of information at the word-, sentence- and text-level. Thus, suggesting that children include constructive processing in their reading goals. Children were most successful at adopting the lexical standard and least successful at adopting the internal consistency standard. This suggests that monitoring difficulties arise when children are required to construct a richly connected situation model, which involves integrating and updating information at the text-level.

This work also explored within error manipulations. The explicitness of information forming the general knowledge violation was manipulated so that information was either related or unrelated to the standard state of affairs in terms of world knowledge. This work extends our understanding of the factors that influence children's monitoring performance. Findings revealed that children use the explicitness of information as a criterion for monitoring their comprehension, particularly when evaluating the sense of information.

The distance between inconsistent information was manipulated so that contradictory information was either adjacent or separated by filler text. There was some evidence that the distance between inconsistent information influenced monitoring. Children made more correct sense judgements when inconsistent information was adjacent than when it was separated by filler text. Also, eye-movements revealed that older children adjusted their reading behaviour depending on the distance between inconsistent information. Further, independent measures analyses revealed that both verbal and numerical working memory were related to monitoring of internal inconsistencies, particularly inconsistencies that were separated by filler text. However, the strength of correlations were weak and no clear pattern of findings were found across experiments. In conjunction, these findings suggest the relationship between comprehension monitoring and working memory capacity is relatively weak. Interestingly, these finding question the importance of working memory capacity as a source of monitoring difficulties. In the context of the situation model, findings suggest that monitoring difficulties may not result from failures in updating information into the situation model. This is because irrespective of the distance between inconsistent information, children were able to update the situation model representation to include inconsistent information. Thus, it seems that monitoring difficulties must arise from an alternative source. One possibility is that children fail to construct a richly elaborate situation model, which includes situation-relevant information that is required to interpret information included later in the text. If children do not build information into their situation model and a later sentence contradicts this information, the problem arises not in updating the situation model and restoring coherence, but in detecting the inconsistent information. Children may encounter difficulties in constructing a richly elaborated situation model either because children struggle with the processing demands associated with the complexity of information or because children do not realise the importance of including situation-relevant information in the model. This finding extends our knowledge of the role of constructing and updating the situation model in the monitoring process, and more broadly reading comprehension.

The real-time data provide insight into children's monitoring strategies. Children spent more time reading error than consistent information. However, perhaps surprisingly this increase in reading time was the only monitoring strategy observed in this work. The eye-movement data revealed that children were highly unlikely to go back and re-read information presented earlier in the text. Thus, it seems that young children in this work primarily monitored their comprehension by undertaking a more purposeful and careful reading of the text.

Furthermore, one area relatively unexplored in monitoring research is the visual presentation of text across the reading surface. This work provides novel insight into this area. The presentation of text did not influence children's ability to evaluate the sense of information or their real-time reading behaviour. This finding suggests that the presentation of text does not impact upon children's ability to integrate information and establish text coherence.

8.4.2. Education

The work included in this thesis has implications for education. Findings revealed that younger children were less successful at monitoring their comprehension than older children. Thus, suggesting that younger children need more support in monitoring their comprehension. One source of support is providing children with instructions that inform them of the need to monitor the sense of information, and encourage the adoption of an 'evaluative mindset'. This work demonstrated that task instructions positively influence monitoring performance.

Consequently, teachers should endeavour to include instructional support reminding children to check their understanding and construct a richly elaborated representation of the situation in classroom instruction. Such instruction may be particularly useful when children are undertaking new or complex tasks, for example monitoring the internal consistency of information. However, instructional support should not be limited to one lesson or particular activities because children need to be able to monitor their comprehension and comprehend information accurately in all aspects of the school curriculum.

Further, the distance between inconsistent information did not strongly influence monitoring behaviour and the relationship between monitoring and working memory was relatively weak. Thus, it is possible that difficulties in monitoring arise from constructing a richly elaborated situation model, rather than updating the situation model. To help children construct a coherent representation of the situation, teachers should include situation model construction as part of the educational methods used in teaching monitoring and reading comprehension. Also, it may be the case that processing demands associated with the complexity of information influence monitoring success. Thus, teachers should strive to provide children with information in its simplest form.

The eye-tracking work revealed that children do not tend to look back and re-read information. Rather, it seems that children prefer to undertake a more careful and purposeful reading of the text. It may be the case that the complexity of information included in passages was not sufficient to compel children to use monitoring strategies beyond spending more time reading the text. Alternatively, it may be the case that young children are not aware of alternative monitoring strategies or have limited experience successfully using alternative monitoring strategies. Thus, it is important that children are taught about different monitoring strategies and remedial action that can be taken to restore comprehension.

8.4.3. Future research

The work included in this thesis has implications for future research. Findings revealed developmental differences in children's monitoring skills. Older children were more likely to correctly judge the sense of information, adjust their reading behaviour in response to error information and successfully detect different types of errors. However, the work only captures the monitoring skills of 7- to 10 year-old children. Thus, further developmental work utilising larger sample sizes with children outside of this age range is required to provide a more comprehensive overview of how comprehension monitoring develops in young children. In addition, longitudinal work may be advantageous because this would allow the developmental time course of monitoring skills to be established.

This work demonstrates that the type of error inserted into texts influences children's comprehension monitoring. Children are better at monitoring their comprehension for nonword than general knowledge violation and internal inconsistency errors. There was also evidence to suggest that within error manipulations such as the explicitness of error information influences children's ability to accurately judge whether or not information that violates general world knowledge makes sense. When creating materials, future experiments that include general knowledge violation errors will need to consider the extent to which error information differs from the standard state of affairs.

In addition, in this work the contradictory information inserted into internal inconsistency passages focused on character or object descriptions. However, it may be the case that different contradictory or passage information, such as that involving temporal (order) aspects, places different demands on children's monitoring skills. To address this, work that manipulates the type of contradictory information inserted into internal inconsistency errors and the content of intervening text is required.

Furthermore, there was some evidence to suggest that the distance between inconsistent information influences monitoring performance. Additional work to explore this finding is required. In Experiments 6-8, contradictory information was separated by three sentences in the far version of passages. Extending the distance between inconsistent information may provide a greater insight into monitoring skill. However, in the literature there is no consensus on the distance between inconsistent information that constitutes a far error. It is likely that different distances will demand different monitoring skills from children. This is because greater distances between inconsistent information will increase the storage and processing demands involved in detecting the error. Thus, it is possible that data will reveal a different pattern of monitoring behaviour and result in different conclusions. This may be particularly likely when making developmental comparisons as children's working memory capacity increases throughout childhood (e.g., Gathercole, 1998). To explore this, future research could be conducted into the distance between inconsistent information.

The work on task instructions suggested that instruction that alerts participants to the likely presence of errors in text change children's reading behaviour. Children's increased reading time in response to this type of instruction may reflect a more purposeful and careful reading of the text, which includes higher levels of evaluation and the adoption of monitoring strategies such as double-checking the sense of information. Thus, future research should carefully consider the nature of task instructions and the influence of this information on monitoring performance, and more generally participant's engagement with the task.

In addition, no consistent pattern of findings was found between comprehension monitoring and word reading, reading comprehension and working memory skills. Future research with a larger sample is necessary to establish the specific skills related to comprehension monitoring, and whether the relative importance of these skills changes during the course of development. Once skills are firmly established, training studies could be conducted in which children are provided with training in the specific skills which are most related to monitoring success.

More generally, this work highlights the advantages of including both off-line and real-time measures to assess comprehension monitoring. These measures tap different aspects of comprehension monitoring and provide different insights into monitoring skills. Thus, to ensure that the multi-dimensional nature of comprehension monitoring is assessed, future experimental work should aim to include both off-line and real-time measures. In addition, the eye-movement experiment provided the opportunity to gain highly naturalistic and fine-grained data. To provide more information of this nature further eye-movement experiments that explore comprehension monitoring developmentally should be undertaken.

8.5. Conclusions

The experimental findings suggest several conclusions. First, in-line with previous research, the data demonstrate developmental differences in comprehension monitoring. Older children were shown to be better at correctly judging the sense of information and more likely to adjust their reading behaviour in relation to error information. It seems that both age groups undertake similar monitoring behaviours, albeit with different levels of success. A range of factors were shown to influence monitoring skill, including the type of error inserted into texts. Children were most able to adopt the lexical standard and least able to adopt the internal consistency standard. There was some evidence that within error manipulations, such as the explicitness of error information and the distance between inconsistent information, influenced children's ability to accurately judge whether or not information makes sense. Also, the extent to which task instructions alerted participants that texts contained errors influenced reading behaviour. Children undertook a more purposeful and careful reading of the text when alerted that texts may contain errors. Furthermore, the data revealed that the relationship between comprehension monitoring and working memory capacity is relatively weak. Thus, suggesting that monitoring difficulties may arise from failures in constructing a richly elaborated situation model, rather than failures in updating the situation model. In the future, longitudinal work that includes a broader age range and larger sample sizes is needed to extend findings and provide more precise information on the developmental trajectory of comprehension monitoring, and the nature of the relationship between monitoring ability and critical component skills.

REFERENCES

Ackerman, B.P. (1984a). The effects of storage and processing complexity on comprehension repair in children and adults. *Journal of Experimental Child Psychology, 37,* 303-334.

Ackerman, B.P. (1984b). Storage and processing constraints on integrating story information in children and adults. *Journal of Experimental Child Psychology, 38,* 64-92.

Adams, M.J. (1994). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.

Albrecht, J.E. & O'Brien, E.J. (1993). Updating a mental model – Maintaining both local and global coherence. *Journal of Experimental Psychology – Learning Memory and Cognition, 19,* 1061-1070.

August, D.L., Flavell, J.H. & Clift, R. (1984). Comparison of comprehension monitoring of skilled and less skilled readers. *Reading Research Quarterly, 20(1),* 39-53.

Baddeley, A.D. & Hitch, G.J. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory (Vol. 8).* New York: Academic Press.

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behaviour, 6(4),* 365-374.

Baker, L. (1984a). Spontaneous versus instructed use of multiple standards for evaluating comprehension: Effects of age, reading proficiency, and type of standard. *Journal of Experimental Child Psychology, 38,* 289-311.

Baker, L. (1984b). Children's effective use of multiple standards for evaluating their comprehension. *Journal of Educational Psychology, 76(4),* 588-597.

Baker, L. (1985). How do we know when we don't understand? Standards for evaluating comprehension. In D.L. Forrest, G.E. MacKinnon & T.G. Waller (Ed.), *Meta-cognition, cognition, and human performance*. New York: Longman.

Baker, L. & Anderson, R.I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly, 17(2),* 281-294.

Baker, L. & Brown, A.L. (1984). Metacognitive skills and reading. In P.J. Pearson, M. Kamil, M.R. Barr & P. Mosenthal (Ed.), *Handbook of reading research*. New York: Longman.

Boyle, J. & Fisher, S. (2008). *Educational testing: A competence-based approach.* London: Wiley.

Breakwell, G.M., Hammond, S., Fife-Schaw, C. & Smith, J.A. (2006). *Research methods in psychology (Third edition).* London: Sage.

Brown, A.L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In P.E. Weinert & R.H. Kluwe (Ed.), *Metacognition, motivation, and understanding*. Hillsdale, NJ: Erlbaum.

Brown, A., Bransford, J.D., Ferrara, R. & Campione, J.C. (1983). Learning, remembering and understanding. In P.H. Mussen, J.H. Flavell & E.M. Markman (Ed.), *Handbook of child psychology (Vol. 3): Cognitive development.* New York: Wiley.

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

Canney, G. & Winograd, P. (1979). *Schemata for reading and reading comprehension performance (Technical Report No. 120).* Urbana: University of Illinois.

Carpenter, P. & Daneman, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixations. *Journal of Verbal Learning and Verbal Behaviour, 20,* 137-160.

Cohen, J., Cohen, P., West, S. & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences (Third edition).* Hillsdale: Erlbaum.

Conners, F.A. (2009). Attentional control and the simple view of reading. *Reading and Writing, 22,* 591-613.

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 and Review, 12 (5),* 769-786.

Cook, A.E. & O'Brien, E.J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes, 51,* 26-49.

Cross, D.R. & Paris, S.G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology, 80,* 131-142.

Daneman, M. & Carpenter, P.A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9,* 561-583.

Ehrlich, M. (1996). Metacognitive monitoring in the processing of anaphoric devices in skilled and less-skilled comprehenders. In C. Cornoldi & J.V. Oakhill (Eds.), *Reading comprehension difficulties: Processes and remediation*. Mahwah, NJ: Erlbaum.

Ericsson, K.A. & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211-245.

Ferretti, T.R., Singer, M. & Patterson, C. (2008). Electrophysiological evidence for the time-course of verifying text ideas. *Cognition, 108,* 881-888.

Field, A. (2005). *Discovering statistics using SPSS (Third edition).* London: Sage.

Foster, H. (2007). *Single word reading test (SWRT).* London: GL Assessments.

Garner, R. (1980). Monitoring of understanding: An investigation of good and poor readers' awareness of induced miscomprehension of text. *Journal of Reading Behaviour, 12(1),* 55-62.

Garner, R. (1981). Monitoring of passage inconsistency among poor comprehenders: A preliminary test of the "piecemeal processing" explanation. *Journal of Educational Research, 74(3),* 159-162.

Garner, R. & Anderson, J. (1982). Monitoring-of-understanding research: Inquiry directions, methodological dilemmas. *Journal of Experimental Education, 50,* 70-76.

Gathercole, S.E. (1998). The development of memory. *Journal of Psychology and Psychiatry, 39,* 3-27.

Gathercole, S.E. & Baddeley, A.D. (1993). *Working memory and language*. East Sussex, UK: BPCC Wheatons.

Gernsbacher, M.A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Erlbaum.

Gilbert, D.T. (1991). How mental systems believe. *American Psychologist, 46,* 107-119.

Glanzer, M., Dorfman, D. & Kaplan, B. (1981). Short-term storage in the processing of text. *Journal of Verbal Learning and Verbal Behaviour, 20,* 656-670.

Glenberg, A.M., Wilkinson, A.C. & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition, 10,* 597-602.

Grabe, M., Antes, J., Thorson, I. & Kahn, H. (1987). Eye fixation patterns during informed and uniformed comprehension monitoring. *Journal of Reading Behaviour, 14(2),* 123-140.

Graesser, A., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative comprehension. *Psychological Review, 101,* 371-395.

Hacker, D.J. (1998). Self-regulated comprehension during normal reading. In D.J. Hacker, J. Dunlosky & A.C. Graesser (Ed.), *Metacognition in educational theory and practice.* Mahwah, NJ: LEA.

Harris, P.L., Kruithof, A., Terwogt, M.M. & Visser, T. (1981). Children's detection of awareness of textual anomaly. *Journal of Experimental Child Psychology, 31*, 212-230.

Herbert, C. & Kubler, A. (2011). Dogs cannot bark: Event-related brain responses to true and false negated statements as indicators of higher-order conscious processing. *PloS One, 6,* e25574.

Isberner, M.B. & Richter, T. (2013). Can readers ignore implausibility? Evidence for nonstrategic monitoring of event-based plausibility in language comprehension. *Acta Psychologica, 142,* 15-22.

Joseph, H.S.S.L., Liversedge, S.P., Blythe, H.I., White, S.J., Gathercole, S.E. & Rayner, K. (2008). Children's and adult's processing of anomaly and implausibility during reading: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology, 61(5),* 708-723.

Juel, C., Griffith, P.L. & Gough, P.B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*, 243-255.

Just, M.A., Carpenter, P.A. & Woolley, J.D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228-238.

Kendeou, P. (2014). Validation and comprehension: An integrated overview. *Discourse Processes, 51,* 189-200.

Kincaid, J.P., Fishburne, R.P., Rogers, R.L. & Chissom, B.S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. *Research Branch Report,* 8-75.

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

Kintsch, W. & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363-394.

Markman, E.M. (1976). Children's difficulty with word-referent differentiation. *Child Development, 47,* 742-749.

Markman, E.M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development, 48(3),* 986-992.

Markman, E.M. (1978). Empirical versus logical solutions to part-whole comparison problems concerning classes and collections. *Child Development, 49,* 168-177.

Markman, E.M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development, 50(3),* 643-655.

Markman, E.M. & Gorin, L. (1981). Children's ability to adjust their standards for evaluating comprehension. *Journal of Educational Psychology, 73,* 320-325.

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

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

McNamara, D.S., Louwerse, M.M. & Graesser, A.C. (2002). Coh-metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension. *Unpublished Grant Proposal, University of Memphis, Memphis, Tennessee.*

Myers, M. & Paris, S.G. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology, 70,* 680-690.

Nation, K., Adams, J.W., Bowyer-Crane, C.A. & Snowling, M. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73(2),* 139-158.

Nation, K. & Snowling, M.J. (1998). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development, 69(4),* 96-1011.

Neale, M.D. (1999). *Neale analysis of reading ability (NARA).* Camberwell, AU: ACER Press.

Nisbett, R.E. & Wilson, T.D. (1977). Telling more than we know: Verbal reports of mental processes. *Psychological Review, 34,* 231-259.

Oakhill, J.V. (1982). Constructive processes in skilled and less-skilled comprehenders' memory for sentences. *British Journal of Psychology, 73,* 13-20.

Oakhill, J.V. (1996). Mental models in children's text comprehension. In J.V. Oakhill & A. Garnham (Eds.), *Mental models in cognitive science.* Hove, UK: Psychology Press.

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

Oakhill, J.V., Hartt, J. & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing, 18*, 657-686.

Oakhill, J.V., Yuill, N. & Garnham, A. (2011). The differential relations between verbal, numerical and spatial working memory abilities and children's reading comprehension. *International Electronic Journal of Elementary Education, 4(1),* 83-106.

Osherson, D. & Markman, E. (1975). Language and the ability to evaluate contradictions and tautologies. *Cognition, 3,* 213-226.

Owings, R.A., Petersen, G.A., Bransford, J.D., Morris, C.D. & Stein, B.S. (1980). Spontaneous monitoring and regulation of learning: A comparison of successful and less successful fifth graders. *Journal of Educational Psychology, 72(2),* 250-256.

Paris, S.G. & Meyers, M. (1981). Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Reading Behaviour, 13*, 5-22.

Patterson, C.J., Cosgrove, J.M. & O'Brien, R.G. (1980). Nonverbal indicants of comprehension and noncomprehension in children. *Developmental Psychology, 16,* 38-48.

Perfetti, C.A. (1985). *Reading ability*. Oxford, England: Oxford University Press.

Perfetti, C.A., Marron, M.A. & Foltz, P.W. (1996). Sources of comprehension failure: Theoretical perspectives and case studies. In C. Cornoldi & J.V. Oakhill (Eds.), *Reading comprehension difficulties: Processes and remediation*. Mahwah, NJ: Erlbaum.

Perfetti, C.A. & McCutchen, D. (1987). Schooled language competence: Linguistic abilities in reading and writing. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics: Reading, writing, and language learning (Vol. 2).* New York: Cambridge University.

Pickering, S. & Gathercole, S.E. (2001). *Working memory test battery for children (WMTB-C).* London: Pearson Assessment.

Poynor, D.V. & Morris, R.K. (2003). Inferred goals in narratives: Evidence from self-paced reading, recall and eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,* 3-9.

Pratt, M. & Nesdale, A.R. (1980). Children's awareness of communication. *Education Research and Perspectives, 7*, 57-68.

Psychology Software Tools, Inc. (E-Prime). (2012). Retrieved from http://www.pstnet.com.

Radvansky, G.A. & Copeland, D.E. (2001). Working memory and situational model updating. *Memory and Cognition, 29,* 1073-1080.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124(3),* 372-422.

Rayner, K., Pollatsek, A., Ashby, J. & Clifton, C. (2012). *Psychology of reading (Second Edition).* New York: Psychology Press.

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

Rinck, M., Gamez, E., Diaz, J.M. & De Vega, M. (2003). Processing of temporal information: Evidence from eye movements. *Memory & Cognition, 31,* 77-86.

Rubman, C.N. & Waters, H.S. (2000). A, B seeing: The role of constructive processes in children's comprehension monitoring. *Journal of Educational Psychology, 92,* 503-514.

Ruffman, T. (1996). Reassessing children's comprehension monitoring skills. In C. Cornoldi & J.V. Oakhill (Eds.), *Reading comprehension difficulties: Processes and remediation*. Mahwah, NJ: Erlbaum.

Schmidt, C.R., Schmidt, S.R. & Tomalis, S.M. (1984). Children's constructive processing and monitoring of stories containing anomalous information. *Child Development, 55,* 2056-2071.

Seigneuric, A., Ehrlich, M., Oakhill, J. & Yuill, N.M. (2000). Working memory resources and children's reading comprehension. *Reading and writing: An interdisciplinary Journal, 13,* 81-103.

Singer, M. (2006). Verification of text ideas during reading. *Journal of Memory and Language, 54,* 574-591.

Singer, M. (2013). Validation in Reading Comprehension. *Current Directions in Psychological Science, 22(5),* 361-366.

Snowling, M.J., Stothard, S.E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., Nation, K. & Hulme, C. (2009). *York assessment of reading comprehension (YARC) – Passage reading.* London: GL Assessment.

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

Towse, J.N., Hitch, G.J. & Hutton, U. (1998). A re-evaluation of working memory capacity in children. *Journal of Memory and Language, 39(2),* 195-217.

Trabasso, T., van den Broek, PW. & Suh, S.Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes, 12,* 1-25.

Tumner, W.E., Nesdale, A.R. & Pratt, C. (1983). The development of young children's awareness of logical inconsistencies. *Journal of Experimental Child Psychology, 36,* 97-108.

Tzeng, Y., van den Broek, P., Kendeou, P. & Lee, C. (2005). The computational implementation of the landscape model: Modeling inferential processes and memory representation of text comprehension. *Behaviour Research Methods, 37,* 277-286.

van den Broek, P., Risden, K. & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inference during reading. In R.F. Lorch & E.J. O'Brien (Ed.), *Sources of coherence in reading.* Hillsdale, NJ: Erlbaum.

van der Schoot, M., Reijntjes, A. & van Lieshout, E.C.D.M. (2011). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing, 25(7),* 1665-1690.

Vosniadou, S., Pearson, P.D. & Rogers, T. (1988). What causes children's failures to detect inconsistencies in text – Representation versus comparison difficulties. *Journal of Educational Psychology, 80,* 27-39.

Wagoner, S.A. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly, 18(3),* 328-346.

Was, C.A. & Woltz, D.J. (2007). Reexamining the relationship between working memory and comprehension: The role of available long-term memory. *Journal of Memory and Language, 56(1),* 86-102.

Wassenburg, S.I., Beker, K., van den Broek, P. & van der Schoot, M. (2015). Children's comprehension monitoring of multiple situational dimensions of a narrative. *Reading and Writing, 28(8),* 1203-1232.

Winograd, P. & Johnston, P. (1982). Comprehension monitoring and the error detection paradigm. *Journal of Reading Behaviour, 14,* 61-76.

Wiswede, D., Koranyi, N., Muller, F., Langner, O. & Rothermund, K. (2012). Validating the truth of propositions: Behavioural and ERP indicators of truth evaluation processes. *Social Cognitive and Affective Neuroscience, 8,* 647-653.

Yuill, N. & Oakhill, J. (1991). *Children's problems in text comprehension.* Cambridge, England: Cambridge Press University.

Yuill, N., Oakhill, J. & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology, 80,* 351-361.

Zabrucky, K. & Moore, D. (1989). Children's ability to use three standards to evaluate their comprehension of text. *Reading Research Quarterly, 24,* 336-352.

Zabrucky, K., Moore, D. & Ratner, H.H. (1985). Comprehension monitoring in good and poor readers: Evidence from verbal and performance measures. *Reading Psychology, 6,* 241-250.

Zabrucky, K. & Ratner, H.H. (1986). Children's comprehension monitoring and recall of inconsistent stories. *Child Development, 57*, 1401-1418.

Zwaan, R.A., Langston, M.C. & Graesser, A.C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6,* 292-297.

Zwaan, R.A. & Radvansky, G.A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123(2),* 162-185.

APPENDICES

APPENDIX 1.1: Passages used in Experiment 1

Nonword
Practice passage
Ben was playing a board game with his sister Ruby.
Ruby rolled the gife and moved her counter forward six places.
Now she was in front of Ben.
Ben was not happy at the thought of losing.
He was going to try very hard to win the game and beat Ruby.

Experimental passages
Yesterday Mike played a tennis match.
He hit every danl with great skill, so won the match.
He was given a gold trophy for doing so well.
Mum was very proud and she bought Mike an ice cream.
Mike is looking forward to his next tennis match.

Ellie's Dad loved to grow plants in his garden.
It was a warm summers day so Ellie used the moxe to water the plants.
Dad was happy that Ellie was helping him.
Dad was worried that the plants might die without the water.
Over the next few days the plants grew very tall.

It was a beautiful summers day.
Rory was playing outside in the gatgen with his sister Clare.
They were playing a game of hide and seek.
Rory closed his eyes and counted to ten before going to look for Clare.
He found her hiding behind a tree.

Yesterday, Chris had a very busy day at work.
He got into his wal to drive home.
It was late and there were only a few cars on the road.
Chris was home in five minutes.
He was tired so ate quickly and went to bed.

Yesterday, Alison was helping her Dad build her new dolls house.
Alison held the wood in place while Dad hit the nail with the happer.
Then they painted the house pink and white.
It took the whole day to build the dolls house.
It looked great when they had finished.

Ruth went to the shop to buy her Dad a present.
She picked a pair of black and white striped football boots.
She took out her rutse and paid the shop owner.
Ruth's Dad was very happy with his new boots.
He wore them to play in Saturday's football match.

Once there was a cat named Tigger.
He had orange and black fur that was very soft.
Tigger had small ears and a beautiful long yaiq.
He liked to play with toy mice and string.
In the evening Tigger liked to spend time by the warm fire.

It was a cold day and the sky was full of clouds.
Vic was walking home when it started to rain.
She opened her umbtenla to keep her dry.
Vic walked quickly and arrived home a few minutes later.
Inside the house it was nice and dry.

Anna and her friends were going out for lunch.
Anna was hungry so chose a steak and chips.
When the food came to the table Anna picked up her knife and fotm to start eating.
The steak was cooked perfectly.
Anna and her friends enjoyed their lunch.

Kim likes to bake cakes.
Every Saturday Kim and her Dad bake a sponge cake.
Kim mixes the flour, milk and eggs in the boyn.
Then Dad puts the mixture into the oven to bake.
Kim is only six years old, so she is not allowed to touch the oven.

It was a cold and icy day.
William was walking home from school when he saw two cars crash.
He opened his bag and called for help on his modine.
Soon he could see some flashing lights and help arrived.
Luckily no one was hurt.

Yesterday the sky was blue and the sun was shining.
Sue went with her friends to the park near her house.
They were going to spend some time in the sun and have a picnic.
In her bavmet she had lots of chocolate cakes.
Everyone enjoyed Sue's cakes.

Eva is six years old and likes to play football.
Her Dad thinks she is a very good player.
Every weekend they practice football skills in the garden.
Eva can kick the bapp very hard.
Once she made Dad cry when she broke a window.

Yesterday Ian had a maths test.
He was worried as maths was not his best subject.
The teacher said the test would take ten minutes.
Ian picked up his penfin to start the first question.
When the test had finished he felt happy because he had done his best.

It was the weekend and Bob was excited to go to his friend's birthday party.
Millie was seven years old.
Bob had bought Millie a new skipping rope.
He had wrapped the rtesent in pink paper and put a bow on top.
Millie loved her new skipping rope.

Anne's Mum and Dad were going away for the weekend.
So she was going to spend a few days at Grandma and Grandpa's house.
Anne was very excited.
She packed her clothes in her suiwfase.
Mum and Dad drove Anne to Grandma and Grandpa's house.

General knowledge violation
Practice passage
Ashley was not feeling very well.
She was very hot and felt sick.
Her Dad was worried so he took her to the zoo to see the doctor.
The doctor gave Ashley some tablets and told her to stay in bed.
A few days later Ashley felt much better.

Experimental passages
This morning Lucy was getting ready for school.
She put her shirt and trousers on, and her shoes on her hands.
Dad drove Lucy to school in the car.
She was early and met her friend in the playground.
They played a skipping game before school started.

Every weekend Phil goes for a bike ride with his Mum.
They sit on their horses and peddle to the park.
They ride around the track three times before going home.
Phil always tries to peddle very fast.
He thinks that he goes faster than Mum.

Bradley likes to read.
Once a week his Mum takes him to the hospital to pick out a new book.
Bradley's favourite books are about cars.
He enjoys looking at the pictures of the cars.
When Bradley is older he wants to drive a fast red car.

Every morning Jack chops wood for his family.
He always uses a knife to chop the wood.
Jack has to chop at least three logs so that there is enough wood for the fire to burn all day.
On school days Jack has to chop the wood quickly.
Otherwise he will be late for school.

Everyday Clare takes the bus to school.
She sits at the front of the bus next to her friend Marcus.
The bus flies through the sky up into the clouds.
The bus is never late.
Clare always arrives at school in time for her first class.

Jenny loved to eat red apples.
She bought an apple tree to grow in her garden.
She dug a big hole with a spoon.
Then she planted the tree and gave it lots of water.
The next year, red apples grew on Jenny's apple tree.

Every morning Peter drinks a large cup of tea when he gets up.
He likes his tea very sweet.
He always stirs in lots of sugar with a fork.
After his tea, Peter gets ready for work.
He drives to work in his sports car.

Once there was a big white rabbit called Bugs.
He liked to play in the farmer's field.
He really liked the ice cream that grew there.
Bugs was very lucky that the farmer had not caught him.
Everyday Bugs ate and played in the farmer's field.

Today, Alf was in his English class.
The teacher told him to write a story about his trip to the seaside.
Alf picked up his toothbrush and started writing with it.
His story was the best in the class.
Alf got a gold star for his story.

It was night time and Amy was getting ready to go to bed.
She changed her clothes.
Then she brushed her teeth with her pen.
Dad read Amy a story about a trip to the beach.
Amy liked the story and asked Dad if they could go to the beach soon.

Yesterday, Kate and her family went out for dinner.
Kate always orders the same thing.
She had a big plate of fish and chips.
The food came to the table and Kate cut into her fish with her scissors.
She told Dad that it tasted very nice.

Roxy and her Dad live on a farm.
Yesterday, Roxy and her Dad had to repair a hole in the fence.
Roxy held the fence in place.
Her Dad hit the nail into the fence with a screwdriver.
When the job was finished the fence looked as good as new.

At school Hayley likes art class.
Today she painted a very pretty picture.
It was of her cat Felix playing with his toy mouse.
She dipped her brush in the gravy and started painting.
She took the picture home to show her Mum and Dad.

Yesterday, Sophie made an apple pie.
She sliced the apples and put them in a glass bowl.
Then she rolled some pastry to go on top of the pie.
Sophie put the pie into the freezer to cook for ten minutes.
The pie tasted so good that Dad had a second slice.

Today Pippa went to the park to play on the swings.
But, she met her friends and they spent all day playing hide and seek.
When she got home she was really thirsty.
She went to the fridge and poured herself a glass of hot juice.
It tasted very good.

Dan's friend Cindy came to his house for lunch.
After lunch Dan washed the dirty plates.
First he stacked the plates.
Then he put the plates in a bowl of mud and washed them until they were clean.
Mum was very happy with Dan for doing such a good job.

Internal inconsistency
Practice passage
John likes cars.
Last week, John bought a new shiny red car.
He spends a lot of time driving with his friends in his new car.
When John drives his new car everyone looks at the lovely green paint.
John is very happy with his new car.

Experimental passages
Emma walked to the shop to buy some sweets.
On her way she lost her purse.
Whilst walking she stopped to talk to her friend Mark.
When she got to the shop, she took out her purse and bought her favourite sweets.
Emma enjoyed eating the sweets on her way home.

At school Sarah likes to play skipping games with her friends.
Yesterday, she broke her leg.
Sarah is seven years old and very tall for her age.
Today at break time she played a skipping game with her friends.
Then the bell went and it was time for class.

Once there was a dog named Max.
He had dark brown fur that was very soft.
Max had a beautiful long tail and big ears.
When he went to the park, all the other dogs wished they had his snow white fur.
Everyday Max was taken for a long walk by his owners.

Last night Hannah walked home through the park.
There was no moonlight, so Hannah could hardly see her way.
Hannah often takes this route home, because it is the quickest way home.
The moon was so bright that it lit the way.
Hannah lives on the other side of the park.

It was the school sports day and Ant was in the running race.
Ant was much slower than all of the other runners.
They all had to run three laps around the track.
Ant easily won the race.
After the race, the runners were very tired.

Tom enjoys flying his kite.
Yesterday, his kite crash landed and broke.
Tom first learnt to fly a kite when he was on holiday with Mum.
Today, Tom flew his kite all day in the park.
Then Mum came to pick him up, so that he could go home for a bath and his dinner.

Ben enjoys ringing his friends on his phone.
Yesterday, his phone fell on the floor and broke.
Ben is a very popular boy, so he has lots of friends.
Today, Ben used his phone to ring his friend Harriet.
He told Harriet about his trip to a museum.

Jack has a grey rabbit named Bugs.
Bugs lives in his cage in the house and never goes outside.
Jack feeds Bugs rabbit food and a carrot every Friday for a treat.
Everyday Bugs plays in the garden on the grass.
Jack really likes Bugs.

Jay has just started at a new school.
Everyday, Mum picks him up and drives him home in the car.
Jay likes to tell Mum what he has done that day.
Jay is always tired after his long walk home.
He has a break then does his homework.

Everyday, Jim walks to school with his big sister Beth.
Jim is five years old and very short.
When they leave for school Beth jumps over the garden gate.
Jim can easily jump over the gate as well, because he is so tall.
They always get to school in time for the bell.

Yesterday, Zoe had a spelling test at school.
The test was very hard and Zoe didn't even try to spell some of the words.
Today, Zoe's teacher gave out the marks from the spelling test.
Zoe got the highest mark because the test was so easy.
Later in the day Zoe had art class.

Jake is a cat that likes chasing mice.
He is a very fat cat.
Yesterday, he chased a mouse that went through a tiny hole in the fence.
Jake followed the mouse through the hole, because he was so thin.
Often Jake does not catch the mice.

Once there was a tiger named Terry.
He was the only tiger that lived on a small island.
Terry always caught his prey, so was a good hunter.
Every night he would hunt with a big group of tigers.
Terry would feed on wild pigs and deer.

Lee has a pet monkey named Mia.
Mia lives in a cage and always sleeps up in the trees.
Lee feeds Mia lots of fruit, she really likes bananas.
At night Mia sleeps on the ground on a bed of leaves.
Lee enjoys watching Mia move about on her hands and feet.

Molly is a small brown mole.
She has very poor eyesight, but her sense of smell is good.
Molly lives under the ground in lots of tunnels.
She easily finds food because she can see so well.
She likes to eat worms, insects and snails.

This morning Mia walked to school.
On her way she lost her pen.
Mia arrived at school in time to see her friends.
In the afternoon, in English class, Mia used her pen to write a story.
She wrote about her fun weekend at the park.

Consistent
Experimental passages
Charlie likes tropical fish.
In his bedroom, he has three fish that live in a big glass tank.
Every morning before school Charlie feeds the fish.
Once a week Charlie helps Mum to clean the tank.
Charlie likes to watch the pretty colours of the fish as they swim.

Chelsea lived on a farm with her Dad.
They had lots of animals, but Chelsea liked the chickens.
On a Monday she would go into the chicken hut to get their eggs.
Her Dad would sell the eggs at the market.
The eggs were popular with people shopping at the market.

Rob bought some sunflower seeds.
He planted the seeds in pots and put them near the window.
When the seeds had grown into plants, Rob put them in the garden and watered them.
The plants grew very tall.
Rob won an award for his tall sunflowers.

Tim likes baking cakes.
On Saturday, he baked a large cake.
He mixed the flour, milk and eggs together in a bowl and put the mixture into the oven.
When the cake was baked, he ate a slice.
He gave the rest to his friends.

Ted has a pet hamster.
It has sandy coloured fur and is named Dora.
Dora was a present for Ted's birthday.
She lives in a large cage in Ted's bedroom.
Once a week Ted takes Dora out of her cage and holds her while Mum cleans the cage.

Rose and her family are going on holiday.
They are going to go on a big plane.
Rose has never flown on a plane before.
She is scared because her brother said that the plane would fly high into the sky.
Dad said Rose should look forward to the adventure.

Today Andy is helping Mum with the weekly food shop.
Mum is reading the list to Andy and he is putting the things into the trolley.
They are going to buy bread, milk and apples.
Andy will carry the bag for Mum.
She thinks he is a big strong boy.

Today, Ellie tripped over a step and dropped a bag of flour on the floor.
She was very sad that she had dropped the flour.
Dad swept up all the flour with a brush.
The floor was clean again.
Ellie will try to be more careful.

It was snowing.
Eve put on her coat and gloves and went outside to build a snowman.
She rolled the snow into a ball to make the body.
She used two stones for eyes and a carrot for a nose.
Dad liked Eve's snowman so much that he took a photo.

After school Stacey loves to play in the park.
She likes the slide the best.
It's fun to climb to the top then come back down to the ground.
Dad always catches her when she comes down the slide.
Today it was raining so they had to go home early.

On Monday morning Joe's Mum baked a big apple pie.
Everyone ate it with custard or ice cream.
They all said it tasted very good.
Joe liked it so much that he asked for a second helping.
Joe was so full that he felt sick.

Dean's Mum is having a baby.
He is excited to have a new baby sister.
Mum says the baby will be small, so Dean will have to be careful.
Dean can't wait to play with his new sister.
He has bought the baby a pink teddy bear.

Meg likes to dance.
After school Dad takes her to a dancing class.
Meg wears a pink dress and shoes.
After class Dad takes Meg for a burger as a treat.
While eating, Meg likes to tell Dad about her class and show him her new dance moves.

Every evening Paul and his Mum look up at the sky.
They like to look at the stars and the moon.
Mum tells Paul the names of the different stars.
Sometimes they will have a treat of cookies and milk.
Paul enjoys spending this special time with Mum.

Yesterday was Matt's birthday.
His Mum had baked him a big chocolate cake.
The cake had icing and sprinkles on top.
Everyone sang Happy Birthday and then Matt blew out the candles.
There were ten candles on the cake because Matt was ten years old.

Dad asked Ivy to help him to paint the fence.
They painted it green.
It took them the whole day to finish the job, but it looked great when they had finished.
Dad was very pleased with the newly painted fence.
He bought Ivy a cake as a treat for her hard work.

Today Joshua is excited because he is going to the seaside.
He is looking forward to playing in the sand and building a sand castle.
Mum has bought him a new bucket and spade for the trip.
Joshua also wants to swim in the sea.
But Mum thinks it will be too cold.

Tomorrow Kirk is going to a water park.
He is excited to go on all the fast slides.
He also wants to swim in the wave pool.
Kirk does not like water in his eyes so will wear his goggles.
Kirk will need lots of energy so is going to bed early.

It was the day of the school fayre.
Richard was helping his teacher on the games stall.
He really likes games where you have to throw a ball through a small hole.
There were lots of great prizes to be won.
Richard's friend Adam won a gold fish.

Once there was an elephant named Nelly.
She lived in a zoo with two elephants.
Nelly had a large trunk, which she used to take in air and suck up water.
She spent most of her day eating fruit.
During the summer she would roll in mud to keep herself cool.

Kay likes to drink cold fruit juice.
Yesterday, when she got home from school, she was thirsty.
Kay went to the fridge and poured herself a large glass of orange juice.
Kay told Dad that it tasted delicious.
So Dad also had a glass of orange juice.

Fay likes to go to the swimming pool.
She loves the warm blue water.
Today was special because Dad had bought Fay a new swimming costume.
It was pink and had big flowers on it.
Fay could not wait to go swimming in her new costume.

Chloe was excited to go to the cinema.
She was going with her friends to watch a new film.
The film was about a boy who had special powers.
Chloe bought a drink and some sweets.
Everyone liked the film and said they wanted to watch it again.

Andrew spilt coffee on the kitchen rug.
He washed the rug and he put it outside to dry.
It was a lovely sunny day and soon the rug was dry.
Andrew put the rug back in the kitchen.
No one found out about him spilling the coffee.

APPENDIX 1.2: Materials used in the background knowledge check for general knowledge violation passages in Experiments 1, 3, 4 and 5

Table 9.1. General knowledge violation background knowledge check questions and response accuracy (%)

Question	Response Accuracy (%)
When you feel sick do you go to the doctors to see the doctor?	97
Do you put your shoes on your feet?	97
When riding a bike do you peddle to make it move?	100
Do you borrow books from the library?	98
Do you use an axe to chop wood?	100
Does a bus drive on the road?	100
When planting a tree in your garden do you use a spade to dig a hole?	97
Do you use a spoon to stir sugar into a cup of tea?	97
Do you keep ice cream in a freezer?	100
Do you write with a pen?	97
Do you brush your teeth with a toothbrush?	100
When eating fish and chips do you use a knife to cut into your fish?	98
Do you use a hammer to hit a nail into a fence?	100
When painting a picture do you dip your brush in paint?	100
When cooking an apple pie do you put it in the oven to cook?	98
Do you pour cold juice from a fridge?	74
After a meal do you wash dirty plates in a bowl of hot soapy water?	79

Note. To control for a yes bias, five questions of a similar structure but not related to the general knowledge violation passages, so are not presented here, were included for which the response was ‘no’. Questions follow passage order shown in Appendix 1.1.

APPENDIX 1.3: Number of passages for each error type within each booklet in Experiment 1

Table 9.2. Number of passages for each error type within each booklet

	Error type		
Booklet	**Nonword**	**General knowledge violation**	**Internal inconsistency**
1	5	6	5
2	5	5	6
3	6	5	5

APPENDIX 2.1: Passages used in Experiment 3

General knowledge violation

Comprehension question instruction
Practice passage
Once there was a big white rabbit called Bugs.
He liked to play in the farmer's field.
He really liked the ice cream that grew there.
Bugs was very lucky that the farmer had not caught him.
Everyday Bugs ate and played in the farmer's field.
Comprehension Question: Was Bugs the name of the rabbit in the story?

Experimental passages
Jenny loved to eat red apples.
She bought an apple tree to grow in her garden.
She dug a big hole with a spoon.
Then she planted the tree and gave it lots of water.
The next year, red apples grew on Jenny's apple tree.
Comprehension Question: Did Jenny plant a green apple tree?

Every morning when Peter gets up he drinks a large cup of tea.
He likes his tea very sweet.
He always stirs in lots of sugar with a fork.
After drinking his tea, Peter gets ready for work.
He drives to work in his sports car.
Comprehension Question: Does Peter drive to work in his sports car?

Yesterday, Jay and his family went out for dinner.
Jay always orders the same thing.
He had a big plate of fish and chips.
The food came to the table and Jay cut into his fish with his scissors.
He told Mum that it tasted very nice.
Comprehension Question: Did Jay eat fish and chips for his dinner?

Roxy and her Dad live on a farm.
Yesterday, Roxy and her Dad had to repair a hole in the fence.
Roxy held the fence in place.
Her Dad hit the nail into the fence with a screwdriver.
When the job was finished the fence looked as good as new.
Comprehension Question: Does Roxy live with her Dad on a farm?

Today Pippa went to the park to play on the swings.
But, she met her friends and they spent all day playing hide and seek.
When she got home she was really thirsty.
She went to the fridge and poured herself a glass of hot juice.
Then she started her homework.
Comprehension Question: Did Pippa spend all day playing in the park with her Mum and Dad?

Yesterday, Sophie made an apple pie.
She sliced the apples and put them in a glass bowl.
Then she rolled some pastry to go on top of the pie.
Sophie put the pie into the freezer to cook for ten minutes.
The pie tasted so good that Dad had a second slice.
Comprehension Question: Is Sophie the name of the main character in the story?

Sense question instruction
Practice passage
This morning Lucy was getting ready for school.
Dad was going to drive Lucy to school.
Before they left the house, Dad brushed Lucy's hair while she put her shoes on her hands.
Lucy met her friends in the playground.
They played a skipping game before the bell rang.

Experimental passages
Bradley likes to read.
Once a week his Mum takes him to the hospital to pick out a new book.
Bradley's favourite books are about cars.
He enjoys looking at the pictures of cars.
When Bradley is older he wants to drive a fast red car.

Every morning Jack chops wood for his family.
He always uses a knife to chop the wood.
Jack has to chop at least three logs so that there is enough wood for the fire to burn all day.
On school days Jack has to chop the wood quickly.
Otherwise he will be late for school.

Today, Joel was in his English class.
The teacher told him to write a story about his trip to the seaside.
Joel picked up his toothbrush and started writing with it.
His story was the best in the class.
Joel got a gold star for his story.

Everyday Daisy takes the bus to school.
She sits at the front of the bus next to her friend Marcus.
The bus flies through the sky up into the clouds.
The bus is never late.
Daisy always arrives at school in time for her first class.

It was night time and Amy was getting ready to go to bed.
She changed her clothes.
Then she brushed her teeth with her pen.
Dad read Amy a story about a trip to the beach.
Amy liked the story and asked Dad if they could go to the beach soon.

At school Lucy likes art class.
Today she painted a very pretty picture.
It was of her cat Felix playing with his toy mouse.
She dipped her brush in the gravy and started painting.
She took the picture home to show her Mum and Dad.

Internal inconsistency

Comprehension question instruction
Practice passage
At school Sarah likes to play skipping games with her friends.
Yesterday, she broke her leg.
Sarah is seven years old and very tall for her age.
Today at break time she played a skipping game with her friends.
Then the bell rang and it was time for class.
Comprehension Question: Is Sarah eight years old?

Experimental passages
Emma walked to the shop to buy some sweets.
On her way she lost her purse.
While walking Emma stopped to talk to her friend Mark.
When she got to the shop, she took out her purse and bought her favourite sweets.
Emma's favourite sweets are fizzy cola bottles.
Comprehension Question: Are Emma's favourite sweets fizzy cola bottles?

Once there was a dog named Max.
He had dark brown fur that was very soft.
Max had a beautiful long tail and big ears.
When he went to the park, all the other dogs wished they had his snow white fur.
Everyday Max was taken for a long walk by his owners.
Comprehension Question: Did Max have small ears?

Lee has a pet monkey named Mia.
Mia lives in a large cage and always sleeps up in the trees.
Lee feeds Mia lots of fruit, she really likes bananas.
At night Mia sleeps on the ground on a bed of leaves.
Lee's friends think that he is very lucky to have such a cool pet.
Comprehension Question: Does Mia really like to eat apples?

It was the school sports day and Kim was in the running race.
Kim was much slower than all of the other runners.
They all had to run three laps around the track.
Kim easily won the race.
After the race, all of the runners were very tired.
Comprehension Question: In the school sports day, did Kim take part in the running race?

Ben enjoys ringing his friends on his phone.
Yesterday, his phone fell on the floor and broke.
Ben is a very popular boy, so he has lots of friends.
Today, Ben used his phone to ring his friend Harriet.
He told Harriet about his trip to the seaside.
Comprehension Question: Does Ben have many friends?

Once there was a tiger named Jim.
He was the only tiger that lived on a small island.
Jim always caught his prey, so he was a good hunter.
Every night he would go out hunting with a big group of tigers.
Jim would feast on wild animals such as deer.
Comprehension Question: Did Jim go hunting every morning?

Sense question instruction
Practice passage
This morning Mia walked to school.
On her way she lost her pen.
Mia arrived at school in time to see her friends.
In the afternoon, in English class, Mia used her pen to write a story.
She wrote about her fun weekend at the park.

Experimental passages
Jack has a grey rabbit named Bugs.
Bugs lives in his cage in the house and never goes outside.
Jack feeds Bugs rabbit food and a carrot every Friday as a treat.
Everyday Bugs plays in the garden on the grass.
Jack really likes Bugs.

Mya has just started at a new school.
Everyday, Dad picks her up and drives her home in the car.
Mya has been going to her new school for a month, so she has lots of new friends.
Mya is always tired after her long walk home.
She has a break then does her homework.

Yesterday, Zoe had a spelling test at school.
The test was very hard and Zoe didn't even try to spell some of the words.
Today, Zoe's teacher gave out the marks from the spelling test.
Zoe got the highest mark.
Later in the day Zoe had art class.

Molly is a small brown mole.
She has very poor eyesight, but her sense of smell is good.
Molly lives under the ground in lots of tunnels.
She easily finds food because she can see so well.
She likes to eat worms, insects and snails.

William enjoys flying his kite.
Yesterday, his kite crash landed and broke.
William first learnt to fly a kite when he was on holiday with Mum.
Today, William flew his kite all day in the park.
Then Mum came to pick him up, so that he could go home for a bath and his dinner.

Last night Hannah walked home through the park.
There was no moonlight, so Hannah could hardly see her way.
Hannah often takes this route home because it is the quickest way home.
The moon was so bright that it lit the way.
Hannah lives on the other side of the park.

Consistent

Comprehension question instruction
It was snowing.
Alicia put on her coat and gloves and went outside to build a snowman.
She rolled the snow into two balls: one for the head and one for the body.
She used stones for eyes and a carrot for a nose.
Dad liked Alicia's snowman so much that he took a photo.
Comprehension Question: Did Alicia use a carrot to make the snowman's nose?

Tim likes baking cakes.
On Saturday, he baked a large cake.
He mixed flour, milk and eggs together in a bowl and put the mixture into the oven.
When the cake was baked, he ate a slice.
He gave the rest to his friends.
Comprehension Question: Did Tim bake a cake on Sunday?

Faye has a pet hamster.
It has sandy coloured fur and is named Ted.
Ted was a present for Faye's birthday.
He lives in a large cage in Faye's bedroom.
Once a week Faye takes Ted out of his cage and holds him while Mum cleans the cage.
Comprehension Question: Does Ted have white fur?

Yesterday was Matt's birthday.
His Mum had baked him a big chocolate cake.
The cake had icing and sprinkles on top.
Everyone sang happy birthday and then Matt blew out the candles.
There were ten candles on the cake because Matt was ten years old.
Comprehension Question: Did Matt's Grandma bake him a big chocolate cake?

Sense question instruction
Rob bought some sunflower seeds.
He planted the seeds in pots and put them near the window.
When the seeds had grown into plants, Rob put them in the garden and watered them.
The plants grew very tall.
Rob won an award for his tall sunflowers.

Today Ellie tripped over a step and dropped a bag of flour on the floor.
She was very sad that she had dropped the bag of flour.
Dad swept up all the flour with a brush.
The floor was clean again.
Ellie will try to be more careful.

Dad asked Ivy to help him to paint the fence.
They painted it green.
It took them the whole day to finish the job, but it looked great when they had finished.
Dad was pleased with the newly painted fence.
He bought Ivy a cake as a treat for her hard work.

Chloe was excited to go to the cinema.
She was going with her friends to watch a new film.
The film was about a boy with special powers.
Chloe bought a drink and some sweets.
Everyone liked the film and said they wanted to watch it again.

APPENDIX 3.1: Passages used in Experiment 4 and 5

Practice passages
Related version
At school Hayley's favourite class is art.
Today, she painted a very pretty picture.
It was of her cat, Oscar, playing with his toy mouse.
Hayley dipped her brush in the gravy and started painting.
First, she painted Oscar's black fur and green eyes.
Then, she painted Oscar's red collar and silver name tag.
Finally, she added a small grey mouse to the picture.
She took the picture home to show her Mum and Dad.

Unrelated version
Once there was a rabbit called Roger.
Roger was a small white rabbit with black ears.
He lived in a burrow with his family of rabbits.
He liked to play in the farmer's field.
He really liked the ice cream that grew there.
Everyday, Roger ate and played in the farmer's field.
Roger was very lucky that the farmer had not caught him.
If the farmer knew that Roger went in his field he might shoot him.

Consistent version
It was a beautiful summers day.
Sam was playing outside in the garden with his sister, Clare.
They were playing a game of hide and seek.
Sam closed his eyes and counted to ten.
Then, he went to look for Clare.
Sam looked under the slide, behind the bin and in the sand pit.
Finally, he found Clare hiding behind a big tree.
Then, it was Clare's turn to count to ten while Sam hid.

Experimental passages
Note. For the error sentence, the related version is presented first, the unrelated version second and the consistent version third.

Jenny loves to eat red apples.
She bought an apple tree to grow in her garden.
Jenny went outside and dug a big hole with a spoon/whisk/spade.
Then, she planted the tree and gave it lots of water.
The next year, red apples grew on Jenny's apple tree.
Jenny had so many apples that she shared some with her friend, Olivia.
Olivia used the apples to make some apple pies.
Everyone said that the apple pies tasted very good.

When Peter wakes up he eats some cereal.
He likes cereal that is the shape of stars the best.
Peter pours lots of the cereal into his big bucket/big bath/big bowl.
While eating, he watches some television.
He usually watches a fun cartoon about toy cars.
Then, Peter puts on his school uniform.
He walks to school with his Mum, Dad and dog, Spot.
Peter meets his friends at the school gate and they go to class.

Joe lives with his family.
Every morning, Joe chops wood for the fire.
Joe always uses a knife/a fork/an axe to chop the wood for the fire.
He has to chop at least three logs.
Otherwise, the wood will run out, and the fire will stop burning.
On school days, Joe has to chop the wood quickly.
If he does not chop the wood quickly, he will be late for school.
Sometimes, when Joe is running late, Dad drives him to school in the car.

Last night, a noise woke Lucy from her sleep.
She looked out of her window to see what was happening.
Outside, in the dark night sky, Lucy could see the sun shining brightly/In the dark night sky, Lucy could see the big football shining brightly/Outside, in the dark night sky, Lucy could see the moon shining brightly.
But, she could not see what might have caused the noise.
She went into Mum and Dad's bedroom.
Mum said she had not heard a noise.
Dad said that it might have been a nightmare.
He took Lucy back to bed, and read a storybook until she fell asleep.

Connor's hobby is swimming.
Every Monday after school, Nan takes Connor to the pool.
Connor puts on his towel/coat/trunks and jumps into the water.
He has just started to swim without armbands.
Connor likes to swim on his back the best.
But, he is also good at swimming on his front.
When Connor is older he wants to take part in swimming races.
Nan thinks Connor can swim fast, so he will easily win the races.

Mary and her family were going on holiday.
Mary was worried about flying, so she sat next to Dad.
Dad held Mary's hand as the car/chair/plane flew through the air up into the clouds.
They talked about what they might do whilst away.
Dad wanted to go fishing and catch a big fish.
Mum wanted to go shopping and buy a new dress.
Mary wanted to go to the zoo and see lots of animals.
They could not wait for their holiday fun to begin.

It was night time, and Emma was getting ready for bed.
First, Emma relaxed in a hot bubble bath.
Then, she put on her pink pyjamas.
Finally, Emma stood at the sink and brushed her teeth with her hairbrush/sponge/toothbrush.
In bed, Dad read Emma a story about a trip to the beach.
Emma liked hearing about playing in the sand and swimming in the sea.
She asked Dad if they could go to the beach soon.
Dad said they could visit the beach next weekend.

Nicola and her Dad live on a small farm.
Yesterday, Nicola and her Dad had to repair a big hole in a fence.
Nicola held the fence in place.
Dad hit a nail into the fence with a bat/ball/hammer.
Then, they painted the fence dark brown.
When the repair job was finished, the fence looked as good as new.
Nicola likes helping Dad with the different jobs on the farm.
Sometimes, Dad buys Nicola a treat to say thank you for her hard work.

Every Sunday, Alex visits his Nan.
Alex likes visiting his Nan because she cooks him a special meal.
His favourite meal is roast beef with potatoes and peas.
Alex likes it when his Nan pours custard/paint/gravy over the beef.
After they have eaten, they will play a card game.
Alex is very good at playing snap.
Before going home, Alex will wash up the dirty plates.
Alex thinks Sunday is the best day of the week.

At the weekend, Jake helps Mum with the housework.
The first job is washing their dirty clothes.
Jake sorts the clothes into dark and light colours.
Then, Mum washes the dark and light clothes in the big dishwasher/big toaster/washing machine.
Mum asks Jake to put all of his toys back in his toy box.
Finally, they dust the surfaces, and mop the floors.
When the jobs are finished, Mum and Jake go to the shop.
They buy a cake as a treat for their hard work.

Michael is ten years old and loves to play football.
Once a week, Mum takes Michael to football training.
At training, he practices lots of ball skills.
Last week, Michael broke his leg so he had to go to the vets/library/hospital.
While his leg heals he will have to miss training.
Michael is very sad that he cannot play football.
Mum has got him a book on football facts to make him feel better.
Also, Michael is going to watch a football match with Dad.

Last weekend, the weather was sunny and warm.
Ryan and his friend Holly took a trip to the seaside.
First, they walked on the beach.
Ryan liked the feeling of the crabs/tigers/sand in between his toes.
While walking on the beach, Holly picked up some pretty shells.
Later, they went for a swim in the sea.
Before leaving the seaside they ate an ice cream.
Ryan and Holly had lots of fun at the seaside.

Yesterday, Dan and his family went out for a special meal.
They were celebrating Dan's tenth birthday.
When Dan goes out for a meal, he always orders the same thing.
He had a big plate of fish, chips and mushy peas.
The food came to the table and Dan cut into his fish with his saw/spade/knife.
For desert, he ate a slice of apple pie.
He told Mum that all of his food had tasted nice.
Dan had lots of fun at his birthday meal.

Today at school, Rebecca was in her art class.
Art is Rebecca's favourite class, because she is very good at painting.
The teacher told Rebecca to paint a picture of her Mum.
First, she drew the outline of her Mum's face.
Then, Rebecca picked up her mop/hoover/brush and started painting with it.
She painted onto the face the eyes, nose and mouth.
Finally, around the face she painted her Mum's long brown hair.
Rebecca got a gold star for her picture.

Yesterday, Kirsty baked some cookies.
She mixed the flour, eggs and milk in a bowl.
Then, Kirsty added the chocolate chips to the mixture.
She cut out twelve cookies, and put them on a tray.
Kirsty put the cookies into the fire/freezer/oven to bake for twenty minutes.
She shared the cookies with her friend, David.
David thought that the cookies tasted very good.
He asked Kirsty if she would show him how to bake the cookies.

This morning, Charlotte went to the park to play with her friends.
They had lots of fun playing on the swings.
They also found time to play a round of crazy golf.
In the afternoon, Charlotte went home.
Charlotte was thirsty so she poured herself a glass of cold juice from the freezer/oven/fridge.
She sat down and started her homework.
Charlotte had to solve ten maths problems.
She asked Dad to help her with some of the more difficult problems.

Adam is a man that loves to drink tea.
Everyday, he drinks two cups of tea.
Adam drinks a cup of tea in the morning and the evening.
He likes to drink sweet tea from a large cup.
Adam always stirs two spoons of salt/sand/sugar into his tea.
Sometimes, he will have one or two biscuits with his tea.
Adam enjoys dipping the biscuits into his tea.
He thinks that tea is one of his favourite drinks.

Last week, Abbie went to the dentist.
Dad took her because she eats lots of sweets.
He thinks the sugar in the sweets is bad for Abbie’s teeth.
At the dentist, Abbie sat in a special chair.
Then, Abbie opened her mouth and the dentist looked at her throat/toes/teeth.
Luckily, the dentist said that this time everything was okay.
Abbie did not like her visit to the dentist.
In the future, she is going to eat fewer sweets.

APPENDIX 3.2: Materials used in the background knowledge check for general knowledge violation passages in Experiments 4 and 5

Table 9.3. General knowledge violation background knowledge check questions and response accuracy (%)

Question	**Response Accuracy (%)**
Do we pour cereals into a bowl?	100
At night, does the sun shine in the sky?	97
When going swimming do boys wear swimming trunks?	93
Do planes fly in the sky?	100
When hitting a nail into a fence do we use a hammer?	100
When eating a roast dinner do we pour gravy over our food?	97
Do we wash dirty clothes in a washing machine?	97
If we break our leg do we go to the hospital?	100
When walking on a beach do we feel sand in between our toes?	97
When painting a picture do we use a brush to paint with?	100
When baking cookies do we put them in the oven to bake?	100
Do we put sugar in our cup of tea?	100
When we go to the dentist do they look at our teeth?	100

Note. To control for a yes bias, four questions of a similar structure but not related to the general knowledge violation passages, so are not presented here, were included for which the response was ‘no’. Questions follow passage order shown in Appendix 3.1.

APPENDIX 3.3: Materials used in verbal and numerical working memory tasks in Experiments 4 – 8

Verbal working memory task materials
Practice Items

1. I like to eat fish and… chips.
   A dog wags its… tail.

2. Zebras have black and white… stripes.
   Leaves grow on… trees.

3. A puppy is a baby… dog.
   There are seven days in a… week.

Level One: Lowest Memory Load

1. I can see with my… eyes.
   A house is made of… bricks.

2. A spider has eight… legs.
   A king wears a… crown.

3. I can hear with my… ears.
   The number after two is… three.

Level Two: Intermediate – Low Memory Load

1. I sat down on a… chair.
   A pig has a curly… tail.
   A bicycle has two… wheels.

2. I wear socks on my… feet.
   I go to sleep in a… bed.
   Cars drive on the… road.

3. A giraffe has a long… neck.
   The sun is in the… sky.
   A postbox is coloured… red.

Level Three: Intermediate – High Memory Load

1. I waved goodbye with my… hands.
   I ran and won the… race.
   A witch flies on a… broom.
   Grass is the colour… green.

2. A duckling is a baby… duck.
   I wear a hat on my… head.
   I eat dinner with a knife and… fork.
   Ice cream is very… cold.

3. Birds have wings to… fly.
   If it rains I will get… wet.
   A farmer lives on a… farm.
   The magician waved his magic… wand.

Level Four: Highest Memory Load

1. Sharks live in the… sea.
   Rockets fly in outer… space.
   When I am thirsty I have a… drink.
   The football team scored a… goal.
   I used a toothbrush to clean my… teeth.

2. To make a snowman you need… snow.
   I tell the time with a wrist… watch.
   I unlocked the door with a… key.
   You can see the moon at… night.
   Worms live under… ground.

3. A fireman helped put out the… fire.
   Everyday I comb my… hair.
   I skip with a skipping… rope.
   Wool comes from… sheep.
   The opposite of wet is… dry.

Numerical working memory task materials

Practice Items

1. 4 – 8 – 4   7 – 0 – 1

2. 4 – 3 – 9   9 – 1 – 2

3. 4 – 9 – 2   7 – 9 – 6   2 – 6 – 9

4. 7 – 0 – 5   1 – 9 – 7   3 – 1 – 2

5. 3 – 9 – 5   4 – 7 – 8   3 – 2 – 7   1 – 9 – 2

6. 3 – 1 – 7   0 – 9 – 4   9 – 7 – 1   1 – 6 – 3

Level One: Lowest Memory Load

1. 4 – 3 – 8   2 – 7 – 5

2. 1 – 7 – 3   0 – 4 – 8

3. 0 – 3 – 7   9 – 1 – 4

Level Two: Intermediate Memory Load

1. 5 – 2 – 8   4 – 3 -1   4 – 8 – 9

2. 6 – 1 – 2   0 – 2 – 8   7 – 9 – 5

3. 5 – 4 – 2   9 – 6 – 9   8 – 2 – 4

Level Three: Highest Memory Load

1. 4 – 0 – 6   7 – 5 – 8   4 – 6 – 1   9 – 2 – 4

2. 0 – 6 – 9   8 – 0 – 1   0 – 2 – 7   9 – 0 – 2

3. 9 – 6 – 4   0 – 2 – 7   5 – 8 – 6   6 – 9 – 3

APPENDIX 3.4: Passages used in Experiment 5

Note. Only the five passages that were revised are presented, all other passages remain the same as those presented in Appendix 3.1. For the error sentence in the experimental passages the related version is presented first, the unrelated version second and the consistent version third.

Practice passage
It was a beautiful summers day.
Sam was playing outside in the garden with his sister, Clare.
They were playing a game of hide and seek.
Sam closed his eyes and counted to ten.
Then, he went to look for Clare.
Sam looked under the slide, behind the bin and in the sand pit.
Finally, he found her hiding behind a big tree.
Then it was Clare's turn to count to ten while Sam hid.

Experimental passages
Adam is a man that loves to drink tea.
Everyday, he drinks two cups of tea.
Adam drinks a cup of tea in the morning and the evening.
He likes to drink sweet tea from a large cup.
Adam stirs two spoons of salt/sand/sugar into his tea.
Sometimes, he will have one or two biscuits with his tea.
Adam enjoys dipping the biscuits into his tea.
He thinks that tea is one of his favourite drinks.

Mary and her family were going on holiday.
Mary was worried about flying, so sat next to Dad.
Dad held Mary's hand as the car/chair/plane flew through the air up into the clouds.
They talked about what they might do whilst away.
Dad wanted to go fishing and catch a big fish.
Mum wanted to go shopping and buy a new dress.
Mary wanted to go to the zoo and see lots of animals.
They could not wait for their holiday fun to begin!

Last night, a noise woke Lucy from her sleep.
She looked out of her bedroom window.
Outside, in the dark night sky, Lucy could see the sun shining brightly/In the dark night sky, Lucy could see the big football shining brightly/Outside, in the dark night sky, Lucy could see the moon shining brightly.
But, she could not see what might have caused the noise.
She went into Mum and Dad's bedroom.
Mum said she had not heard a noise.
Dad said that it might have been a nightmare.
He took Lucy back to bed, and read a storybook until she fell asleep.

Yesterday, Kirsty baked some cookies.
She mixed the flour, eggs and milk in a bowl.
Kirsty added the chocolate chips to the mixture.
She cut out ten cookies, and put them on a tray.
Kirsty put the cookies into the fire/freezer/oven to bake for twenty minutes.
She shared the cookies with her friend, David.
David thought that the cookies tasted very good.
He asked Kirsty if she would show him how to bake the cookies.

APPENDIX 4.1: Internal inconsistency passages used in Experiment 6

Practice passages

Near version

Emma went to the shop to buy her Dad a present.
She picked a pair of black and white football boots.
She took out her purse and paid the shop owner.
Dad really liked his new red and blue football boots.
He wore them to play in Saturday's football match.
In the match, he scored three goals.
Emma told Dad that the new boots made him play football better.
Dad said the boots were lucky and would wear them in next weeks match.

Far version

Jenny has a pet cat called Felix.
She feeds Felix lots of food, so he is a very fat cat.
Felix has orange fur that is very soft.
He loves to spend time chasing mice.
Yesterday, Felix chased a mouse that went through a tiny hole in the fence.
He followed the mouse through the hole, because he is so thin.
Often, Felix does not catch the mice.
So that he does not get sad, Jenny has bought Felix a toy mouse to play with.

Consistent version

Yesterday, Kirsty and her family went out for dinner.
They were celebrating Kirsty's seventh birthday.
When Kirsty goes out for dinner, she always orders the same thing.
She had a burger and chips.
The food came to the table, and Kirsty could not wait to start eating.
For desert, she ate a large slice of chocolate cake and ice cream.
She told Dad that the cake tasted very nice.
Kirsty had lots of fun at her birthday dinner.

Experimental passages

Note. The far internal inconsistency version of passages is presented. The position of inconsistent information for the near internal inconsistency version is indicated by the asterisk and information for the consistent version is presented in parenthesis.

Emily and her family were going on holiday.
They all got onto the plane and put on their seat belts.
Everyone was excited about spending two weeks in the sun.
*
Dad said the hotel that they were staying in had a pool.
Emily could not wait to go swimming.
*Before they knew it, two hours had passed and the family got off the boat.
(Before they knew it, two hours had passed and the family got off the plane.)
Mum was looking forward to reading in the sun.
Dad hoped that he would find time to play golf.

Sarah was outside playing on her roller skates.
She fell over and hurt her arm.
Sarah had only just started to learn how to skate.
*
Dad told Sarah that she must be more careful when skating.
Sarah said she would be more careful in the future.
*Dad took Sarah to the hospital to have an X-Ray of her leg.
(Dad took Sarah to the hospital to have an X-Ray of her arm.)
Luckily, Sarah had not broken any bones this time.
As a treat Dad bought Sarah an ice cream on the way home.

Once there was a dog named Max.
He was covered from head to toe in dark brown fur.
Max had a beautiful long tail and big floppy ears.
*
Everyday, his owners took him for a long walk.
Max enjoyed going for walks in the fields.
*When he went outside, all of the other dogs wished they had his snow white fur.
(When he went outside, all of the other dogs wished they had his dark brown fur.)
On his walks Max liked to chase other dogs.
Sometimes, his owners would take a ball to play fetch with Max.

Once there was a tiger called Harry.
He was the only tiger that lived on a small island.
(He was the oldest tiger that lived on a small island.)
During the day, to avoid the heat, Harry rested in the grass.
*
Harry always caught the animals that he hunted, so he was a good hunter.
He would hunt wild animals such as deer.
*Every night, he would go out hunting with a big group of tigers.
Sometimes, he would hunt monkeys.
But, he most liked hunting wild pigs.

It was the school sports day and Sophie was in the running race.
Sophie was much slower than all of the other runners.
They all had to run three laps around the track.
*
For the race Sophie wore her new trainers.
She also had a new water bottle to drink from when she got thirsty.
*Sophie easily won the race.
(Sadly, Sophie lost the race.)
After the race Sophie was very tired.
Sophie found her Dad, and he took her home.

Today, Chloe had a spelling test at school.
The test was very hard and Chloe didn't spell any of the words.
Later, Chloe’s teacher would give out the marks for the test.
*
In the afternoon, Chloe was looking forward to art class.
Art is one of Chloe’s favourite classes.
*Chloe got the highest mark in the spelling test.
(Chloe did not do well in the spelling test.)
In art class, she drew her dog, Spot.
She took the drawing home to show her Mum and Dad.

Jessica is a monkey that lives at a zoo.
Jessica always sleeps up in the branches of tall trees.
(During the day, Jessica sits in the branches of trees.)
Her fur is mostly black, but her tail is white.
*
Jessica likes to eat lots of different fruit.
Often, the zookeepers will feed her apples.
*At night, Jessica sleeps on the ground on a bed of leaves.
Sometimes, she will be fed a banana as a treat.
Jessica is one of the zoos most popular animals.

Last week it was James’ birthday.
He was eleven years old.
Mum baked James a big chocolate sponge cake.
*
The cake had lots of icing and sprinkles on top.
At his party, everyone sang happy birthday to James.
*There were twelve candles on the cake; the perfect number.
(There were eleven candles on the cake; the perfect number.)
Then James cut the cake, it tasted so good that he asked Mum for a second slice.
James told his Mum that he had really enjoyed his birthday.

At school, Kate likes to play with her friends in the playground.
Yesterday, she broke her leg and it was put in a cast.
At break time, Kate enjoys playing skipping games.
*
But, Kate’s favourite type of games are bat and ball games.
She has lots of fun when she plays rounders.
*Today, Kate played a skipping game with her friends.
(Today, Kate watched her friends play a skipping game.)
Kate’s teachers think she is very sporty.
When she is older, Kate wants to teach children to play sports.

Amy and Jack walked to the shop to buy some sweets.
On their way Amy lost her purse.
To get to the shop, Amy and Jack walked past their school.
*
Amy's favourite sweets are jelly beans.
She likes the red and green jelly beans the best.
*When they got to the shop, Amy took out her purse and bought some sweets.
(When they got to the shop, Jack took out his money and bought some sweets.)
Jack prefers to eat fizzy cola bottles.
He really enjoys the sour taste of the cola bottles.

Molly is a small brown mole.
She has very poor eyesight, but her sense of smell is good.
Molly lives underground in lots of tunnels.
*
She sleeps in a nest made from grass and leaves.
Her short fur means that she can easily move around the tunnels and her nest.
*She easily finds food because she can see so well.
(She easily finds food with her good sense of smell.)
Molly eats worms, insects and snails.
Sometimes, she will eat small mammals such as mice.

Hannah has just started at a new school.
Everyday, Dad picks her up and drives her home in the car.
Hannah likes her new school, but she misses her old teachers.
*
At school Hannah has made lots of friends, her best friend is Lily.
Hannah and Lily are in the same class.
*After her long walk home, Hannah is always tired.
(After her day at school Hannah is always tired.)
When Hannah gets home she watches television and eats a snack.
Then she meets Lily at the park to play on the swings.

Matthew walks to school with his big sister Lauren.
He is five years old and very short.
When they leave for school Lauren jumps over the garden gate.
*
Lauren is always ready for school.
But, sometimes Matthew has to put his tie on as they leave the house.
*Matthew can easily jump over the gate as well, because he is so tall.
(When Matthew is taller he will also be able to jump over the gate.)
They always get to school in time for the bell.
Lauren puts their lunch boxes in the hall before they go into class.

Today was a special day: it was Mum's birthday.
Oliver went to the shops and bought a bunch of pink flowers for his Mum.
Later in the day, the family were going out for a meal.
*
Mum asked Oliver what he might like to eat.
Oliver said he wanted to eat pasta.
*Oliver gave his Mum the yellow flowers and she put them in a vase.
(Oliver gave his Mum the pink flowers and she put them in a vase.)
Mum thought she would like to eat pizza.
They decided to go out for an Italian meal.

Tom went to watch a circus show.
He was excited as he entered the green and white circus tent.
A lady on very tall sticks took his ticket.
*
Tom hoped to see clowns performing funny tricks.
Also, he wanted to see a lion jump through a flaming hoop.
*The show started and bright lights lit up the red and blue circus tent.
(The show started and bright lights lit up the green and white circus tent.)
Tom really enjoyed the show.
He would like to go again next year.

William enjoys flying his kite in the local park.
Last night, his kite crash landed and broke.
William has been flying kites since he was a little boy.
*
William's first kite was the shape of a spaceship.
He likes flying big kites that are brightly coloured.
*Today, William flew his kite all day in the park.
(Today, William bought a new kite to fly in the park.)
When it began to get dark Mum came to pick him up.
He went home and had a bath before eating his dinner.

Josh has a rabbit named Bugs.
Bugs lives in a cage in the living room, and never goes outside.
(Bugs lives in a cage in the living room, but likes going outside.)
Josh's Mum and Dad bought Bugs as a birthday present for Josh.
*
Bugs has grey and white fur that is very soft.
Josh feeds Bugs rabbit food and water.
*Everyday, Bugs plays in the garden on the grass.
As a special treat, Josh gives Bugs a carrot to eat every Friday.
Josh really likes having Bugs as his pet.

Last night, Ben walked home through the park.
There was no moonlight, so Ben could hardly see his way.
Ben walked through the park because it is the quickest way home.
*
Ben lives on the other side of the park.
It usually takes him ten minutes to walk home.
*The moon was so bright that it lit his way.
(He used a bright torch to help light the way.)
Ben was looking forward to having a hot bath and eating his dinner.
Mum had cooked him burger and chips.

APPENDIX 4.2: Internal inconsistency passages used in Experiment 7

Practice passages

Adjacent version

Emma went to the shop to buy her Dad a present.
Emma picked a pair of black football boots.
Dad really liked his new red football boots.
He wore them to play in Saturday's football match.
In the match, Dad scored three goals.
Emma told Dad that the new boots made him score more goals.
Dad thought that the new boots were lucky.
He told Emma that he would wear them in next weeks match.

Far version

Jenny has a pet cat called Felix.
She feeds Felix lots of food, so Felix is a very fat cat.
Felix has orange fur that is very soft.
He loves to spend time chasing mice.
Yesterday, Felix chased a mouse that went through a hole in a fence.
He followed the mouse through the hole, because he is so thin.
Often, Felix will not catch the mice.
So that Felix does not get sad, Jenny has bought him a toy mouse to play with.

Consistent version

Yesterday, Kirsty and her family went out for dinner.
They were celebrating Kirsty's seventh birthday.
When Kirsty goes out for dinner, she always orders the same thing.
She had a burger and chips.
The food came to the table and Kirsty could not wait to start eating.
For desert, she ate a large slice of chocolate cake and ice cream.
Kirsty told Dad that the cake tasted very nice.
Kirsty had lots of fun at her birthday dinner.

Experimental passages

Note. The far internal inconsistency version of passages is presented. The position of inconsistent information for the adjacent internal inconsistency version is indicated by the first asterisk, the position of the near internal inconsistency version is indicated by the second asterisk and information for the consistent version is presented in parenthesis.

Emily and her family were going on holiday.
They all got onto the plane and put on their seat belts.
*
Everyone was excited about spending two weeks in the sun.
*
Dad said the hotel that they were staying in had a pool.
Emily could not wait to go swimming.
*Before they knew it two hours had passed and the family got off the boat.
(Before they knew it two hours had passed and the family got off the plane.)
Mum was looking forward to reading in the sun.
Dad hoped that he would find time to play golf.

Sarah was outside playing on her roller skates.
Sarah fell over and hurt her arm.
*
Sarah has only just started to learn how to skate.
*
Dad told Sarah that she must be more careful when skating.
Sarah said she would be more careful in the future.
*Dad took Sarah to the hospital to have an X-Ray of her leg.
(Dad took Sarah to the hospital to have an X-Ray of her arm.)
Luckily, Sarah had not broken any bones this time.
As a treat Dad bought Sarah an ice cream on the way home.

Once there was a dog named Max.
Max was covered in dark brown fur.
*
Also, Max had a beautiful long tail and big floppy ears.
*
Everyday his owners took him for a long walk.
Max enjoyed going for walks in the fields.
*When Max went outside, all of the other dogs wished they had his snow white fur.
(When Max went outside, all of the other dogs wished they had his dark brown fur.)
On his walks Max liked to chase other dogs.
Sometimes his owners would take a ball to play fetch with Max.

Once there was a tiger called Harry.
Harry lived on a small island and was the only tiger living there.
(He lived on a small island and was the oldest tiger living there.)
*
Harry had orange and black fur and he was a strong tiger.
*
Harry always caught the animals that he hunted, so he was a good hunter.
He would hunt wild animals such as deer.
*Harry would go out hunting with a big group of tigers every night.
Sometimes, he would hunt monkeys.
But, he most liked hunting wild pigs.

It was the school sports day and Sophie was in the running race.
Sophie was much slower than all of the other runners.
*
Sophie had to run three laps around the track.
*
For the race Sophie wore her new trainers.
She also had a new water bottle to drink from when she got thirsty.
*Sophie easily won the race.
(Sadly, Sophie lost the race.)
After the race Sophie was very tired.
Sophie found her Dad, and he took her home.

Today Chloe had a spelling test at school.
The test was very hard and Chloe didn't spell any of the words.
*
Chloe's class has a new spelling test every week.
*
In the afternoon, Chloe was looking forward to art class.
Art is one of Chloe's favourite classes because she is good at drawing.
*Chloe got the highest mark in the spelling test.
(Chloe did not do well in the spelling test.)
In art class she drew her dog, Spot.
She took the drawing home to show her Mum and Dad.

Jessica is a monkey that lives at a zoo.
Jessica loves to spend her time sleeping, she always sleeps up in the branches of tall trees.
(Jessica loves to spend her time sleeping, during the day she sits in the branches of trees.)
*
Jessica's fur is mostly black, but her tail is white.
*
Jessica likes to eat lots of different fruit.
Often, the zookeepers will feed her apples and oranges.
*At night, Jessica sleeps on the ground on a bed of leaves.
Sometimes, she will be fed a banana as a treat.
Jessica is one of the zoos most popular animals.

Last week it was James' birthday.
James was eleven years old.
*
Mum had baked James a big chocolate sponge cake.
*
The cake had lots of icing and sprinkles on top.
At his party, everyone sang Happy Birthday to James.
*There were twelve candles on James' cake; the perfect number.
(There were eleven candles on James' cake; the perfect number.)
Then James cut the cake, it tasted so good that he asked Mum for a second slice.
James told his Mum that he had really enjoyed his birthday.

At school Kate likes to play with her friends in the playground.
Yesterday, Kate broke her leg and it was put in a cast.
*
At break time, Kate enjoys playing skipping games.
*
But, Kate's favourite type of games are bat and ball games.
She has lots of fun when she plays rounders.
*Today, Kate played a skipping game with her friends.
(Today, Kate watched her friends play a skipping game.)
Kate's teachers think she is very sporty.
When she is older Kate wants to teach children to play sports.

Last weekend, Amy and her friend Jack walked to the shop to buy some sweets.
On their way Amy lost her purse.
*
While walking to the shop, Amy and Jack talked about their favourite sweets.
*
Amy's favourite sweets are jelly beans.
She likes the red, black and green jelly beans the best.
*When they got to the shop, Amy took out her purse and bought some sweets.
(When they got to the shop, Jack took out his money and bought some sweets.)
Jack prefers to eat fizzy cola bottles.
He really enjoys the sour taste of the cola bottles.

Molly is a small brown mole.
Molly has very poor eyesight but her sense of smell is good.
*
Molly lives deep underground in lots of tunnels.
*
She sleeps in a nest made from grass and leaves.
Her short fur means that she can easily move around the tunnels and her nest.
*Molly easily finds food because she can see so well.
(Molly easily finds food with her good sense of smell.)
Molly eats worms, insects and snails.
Sometimes she will eat small mammals such as mice.

Hannah has just started at a new school.
Everyday Dad picks Hannah up and drives her home in the car.
*
Hannah likes her new school, but she misses her old teachers.
*
At school Hannah has made lots of friends, her best friend is Lilly.
Hannah and Lilly are in the same class.
*After Hannah's long walk home she is always tired.
(After Hannah's long day at school she is tired.)
When Hannah gets home she watches television and eats a snack.
Then she meets Lilly at the park to play on the swings.

Matthew walks to school with his big sister Lauren.
Matthew is five years old and very short.
*
When getting ready for school, Matthew struggles to put his tie on.
*
Lauren is always ready for school.
So she is able to help Matthew put his tie on as they walk to school.
*When leaving the house, Matthew jumps over the garden gate because he is so tall.
(When leaving the house, Matthew waves to Mum as he walks down the garden path.)
They arrive at school in time for the bell.
Lauren puts their lunch boxes in the hall before they go into class.

Today was a special day: it was Mum's birthday.
Oliver went to the shops and bought a bunch of pink flowers for his Mum.
*
Later in the day the family were going out for a meal.
*
Mum asked Oliver what he might like to eat.
Oliver said he wanted to eat pasta.
*Oliver gave his Mum the yellow flowers and she put them in a vase.
(Oliver gave his Mum the pink flowers and she put them in a vase.)
Mum thought she would like to eat pizza.
They decided to go out for an Italian meal.

Tom went to watch a circus show.
Tom was very excited as he entered the green and white circus tent.
*
A lady standing on very tall sticks took his ticket.
*
Tom hoped to see clowns performing funny tricks.
Also, he wanted to see a lion jump through a flaming hoop.
*The red and blue circus tent lit up as the show started.
(The green and white circus tent lit up as the show started.)
Tom really enjoyed the show.
He would like to go again next year.

William enjoys flying his kite in the local park.
William was flying his kite last night when it crashed and broke.
*
William has been flying kites since he was a little boy.
*
William's first kite was the shape of a spaceship.
He likes flying big kites that are brightly coloured.
*Today, William flew his kite all day in the park.
(Today, William bought a new kite to fly in the park.)
When it began to get dark Mum came to pick him up.
He went home and had a bath before eating his dinner.

Josh has a rabbit named Bugs.
Bugs lives in a cage in the living room, and never goes outside.
(Bugs lives in a cage in the living room, but likes going outside.)
*
Josh's Mum and Dad bought Bugs as a birthday present for Josh.
*
Bugs has grey and white fur that is very soft.
Josh feeds Bugs rabbit food and water.
*Everyday, Bugs plays in the garden on the grass.
As a special treat, Josh gives Bugs a carrot to eat every Friday.
Josh really likes having Bugs as his pet.

Last night, on his way home from a friends party, Ben walked through the park.
There was no moonlight, so Ben could hardly see his way home.
*
Ben walked through the park because it is the quickest way home.
*
Ben lives on the other side of the park.
It usually takes him ten minutes to walk home.
*The moon was so bright that it lit Ben's way home.
(Ben used a bright torch to help light the way home.)
Ben was looking forward to having a hot bath and eating his dinner.
Mum had cooked him burger and chips.

Nan has moved to a new house.
Nan's favourite room in the new house is her bedroom because it is very big.
*
Nan has painted all of the walls in her new house pink.
*
Nan is happy that the new house has a garden.
She cannot wait to sit outside in the sun.
*Nan's bed did not fit into her new bedroom because the room is very small.
(There is lots of space in Nan's new bedroom because the room is very big.)
Nan loves her new house.
She is having a party to show the house to her family.

Saturdays are Dan's favourite day of the week.
In the morning, Dan will get up early and go on a bike ride with Mum.
*
Whilst they are out they talk about what to eat for breakfast.
*
Every week they enjoy eating a different breakfast.
Dan likes it when they have pancakes.
*Dan and Mum ride their horses to the shops to buy food for their breakfast.
(Dan and Mum ride their bikes to the shops to buy food for their breakfast.)
At the shop, Mum always lets Dan choose what they will buy.
Last week Dan chose bacon and eggs.

This morning Jenny walked to school with her Mum.
On Jenny's way to school she lost her red pen.
*
Jenny had a very busy day at school.
*
Jenny's first lesson was drama.
The class were performing a play about a boy that lived in the jungle.
*When Jenny got to school she used her red pen to write in her diary.
(When Jenny got to school she used her blue pen to write in her diary.)
Later in the day, as part of Jenny's maths lesson, she had a mental maths test.
After school Jenny went to netball practice.

Peter is a five year old boy.
Peter always goes to bed at six o’clock.
*
Before going to bed Peter eats a snack, has a bath and brushes his teeth.
*
Peter likes to wear his spider man pyjamas to bed.
He thinks that they help him to sleep better.
*Last night, Peter went to bed at ten o’clock.
(Last night, Peter went to bed at six o’clock.)
Mum will read Peter a story before turning off the lights.
Then she gives him a kiss and goes downstairs to watch television.

After a busy day at work, Dad goes home to relax.
When Dad gets home he always has a bath.
*
Dad likes using shampoo that smells of apples.
*
Dad also enjoys using a special sponge that helps him to scrub his back.
Sometimes, when Dad is in the bathroom he will play rock music on the radio.
*Last night, when Dad got home he had a shower.
(Last night, when Dad got home he had a bath.)
Dad has just painted the bathroom.
He thinks that the bathroom is his favourite room in the house.

Emma was looking forward to a day at the zoo.
Emma was going on a special family trip with her Mum and Dad.
*
During their visit they saw some monkeys eating lots of bananas.
*
They also saw seals performing tricks.
Emma even got to throw a fish to one of the seals.
*Emma had lots of fun with her class on her school trip to the zoo.
(Emma had fun with her Mum and Dad on their family trip to the zoo.)
Emma would like to visit the zoo again soon.
Next time she wants to see the tigers.

APPENDIX 4.3

Table 9.4. Mean raw scores and standard scores (with standard deviations and t-tests) for word reading, reading comprehension and working memory measures by age group

Measure	**Experiment 6** 7- to 8-year-olds (n = 26)	**Experiment 6** 9- to 10-year-olds (n = 28)	***t*(52)**	***d***	**Experiment 7** 7- to 8-year-olds (n = 21)	**Experiment 7** 9- to 10-year-olds (n = 22)	***t*(41)**	***d***
Word reading								
Raw score (maximum = 104)	61.35 (7.79)	72.50 (5.83)	5.99***	1.62	66.00 (6.99)	69.64 (8.07)	1.58	.48
Standardised score	115.31 (10.28)	112.75 (8.03)	1.02	.28	117.10 (7.13)	102.73 (9.89)	5.44***	1.67
Nonword reading								
Raw score (maximum = 63)	33.77 (10.10)	41.18 (7.17)	3.13**	.85	37.29 (8.96)	42.27 (8.48)	1.88	.57
Standardised score	115.96 (12.07)	113.89 (10.22)	< 1	-	117.52 (10.01)	111.82 (12.10)	1.68	.51
Reading rate								
Raw score (seconds)	92.98 (24.92)	106.18 (21.21)	2.10*	.57	114.07 (27.01)	110.75 (23.07)	< 1	-
Ability score	65.15 (10.38)	79.61 (6.98)	6.04***	1.63	70.05 (7.00)	77.27 (8.58)	3.02**	.92
Standardised score	111.42 (11.42)	112.50 (10.00)	< 1	-	112.57 (7.83)	106.18 (11.76)	2.09*	.64
Word reading accuracy								
Raw score	3.71 (1.90)	3.48 (2.05)	< 1	-	4.95 (2.82)	3.34 (2.27)	2.07	.63
Ability score	53.15 (8.22)	63.79 (7.74)	4.90***	1.33	56.05 (8.11)	63.50 (7.48)	3.13**	.95
Standardised score	110.46 (10.89)	111.39 (10.90)	< 1	-	111.48 (10.68)	108.63 (10.97)	< 1	-
Reading comprehension								
Raw score (maximum = 8)	4.87 (1.11)	5.38 (1.01)	1.76	.48	4.74 (1.26)	4.93 (1.21)	< 1	-
Ability score	54.96 (7.02)	64.57 (6.75)	5.13***	1.40	57.76 (7.34)	61.91 (7.50)	1.83	.56
Standardised score	106.96 (8.69)	109.07 (9.73)	< 1	-	107.62 (9.57)	103.09 (11.07)	1.43	.44
Verbal working memory								
Number of items (maximum = 42)	13.73 (5.66)	19.04 (6.51)	3.19**	.87	11.76 (5.23)	17.27 (6.80)	2.97**	.91
Number of trials (maximum = 12)	3.96 (1.48)	5.25 (1.53)	3.14**	.86	3.62 (1.53)	4.95 (1.53)	2.86**	.87
Span (maximum = 5)	2.35 (.56)	2.71 (.66)	2.20*	.59	2.10 (.62)	2.64 (.58)	2.94**	.90
Numerical working memory								
Number of items (maximum = 27)	12.27 (5.70)	20.50 (5.15)	5.58***	1.52	14.62 (5.57)	16.68 (4.99)	1.28	.39
Number of trials (maximum = 9)	3.77 (1.77)	6.07 (1.68)	4.90***	1.33	4.71 (1.76)	5.14 (1.55)	< 1	-
Span (maximum = 4)	2.19 (.57)	3.18 (.67)	5.82***	1.59	2.48 (.68)	2.68 (.57)	1.08	.32

Note. Raw scores on the upper row, ability scores on the middle row and standardised scores on the lower row. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

APPENDIX 5.1: Internal inconsistency passages used in Experiment 8

Note. Only the seven passages that were revised are presented, all other passages remain the same as those presented in Appendix 4.2. For the experimental passages, the far internal inconsistency version of passages is presented. The position of inconsistent information for the adjacent internal inconsistency version is indicated by the first asterisk, the position of the near internal inconsistency version is indicated by the second asterisk and information for the consistent version is presented in parenthesis.

Practice passage

Consistent version
Yesterday, Kirsty and her family went out for dinner.
They were celebrating Kirsty's seventh birthday.
When Kirsty goes out for dinner, she always orders the same thing.
She had a burger and chips.
The food came to the table, and Kirsty could not wait to start eating.
For desert, she ate a large slice of chocolate cake and ice cream.
Kirsty told Dad that the cake tasted very nice.
Kirsty has lots of fun at her birthday dinner.

Experimental passages
Once there was a tiger called Harry.
Harry lived on a small island, and was the only tiger living there.
(Harry lived on a small island, and was the oldest tiger living there.)
*
Harry had orange and black fur and he was a strong tiger.
*
Harry always caught the animals that he hunted, so he was a good hunter.
He would hunt wild animals such as deer.
*Harry would go out hunting with a big group of tigers every night.
Sometimes, he would hunt monkeys.
But, he most liked hunting wild pigs.

At school, Kate likes to play with her friends in the playground.
Yesterday, Kate broke her leg and it was put in a cast.
*
At break time, Kate enjoys playing skipping games.
*
But, Kate's favourite type of games are bat and ball games.
She has lots of fun when she plays rounders.
*Today, Kate played a skipping game with her friends.
(Today, Kate watched her friends play a skipping game.)
Kate's teachers think she is very sporty.
When she is older Kate wants to teach children to play sports.

Matthew walks to school with his big sister Lauren.
Matthew is five years old and very short.
*
When getting ready for school, Matthew struggles to put on his tie.
*
Lauren is always ready for school.
So she is able to help Matthew put on his tie as they walk to school.
*When leaving the house, Matthew jumps over the garden gate because he is so tall.
(When leaving the house, Matthew waves to Mum as he walks down the garden path.)
They arrive at school in time for the bell.
Lauren puts their lunch boxes in the hall before they go into class.

Tom went to watch a circus show.
Tom was excited as he entered the green and white circus tent.
*
A lady on very tall sticks took his ticket.
*
Tom hoped to see clowns performing funny tricks.
Also, he wanted to see a lion jump through a flaming hoop.
*As the show started bright lights lit up the red and blue circus tent.
(As the show started bright lights lit up the green and white circus tent.)
Tom really enjoyed the show.
He would like to go again next year.

This morning Jenny walked to school.
On Jenny's way to school she lost her red pen.
*
Jenny had a very busy day at school.
*
Jenny's first lesson was drama.
The class were performing a play about a boy that lived in the jungle.
*When Jenny got to school she used her red pen to write in her diary.
(When Jenny got to school she used her blue pen to write in her diary.)
Later in the day, as part of Jenny's maths lesson she had a mental maths test.
After school Jenny went to netball practice.

After a busy day at work, Dad goes home to relax.
When Dad gets home he always has a bath.
*
Dad likes using shampoo that smells of apples.
*
Dad also enjoys using a special sponge that helps him to scrub his back.
Sometimes when Dad is in the bathroom he will play rock music on the radio.
*Last night, when Dad got home he had a shower.
(Last night, when Dad got home he had a bath.)
Dad has just painted the bathroom.
He thinks that the bathroom is his favourite room in the house.