



A University of Sussex DPhil thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

**Reading Comprehension in Adults:
Component Skills; False Memories; and Judgements
of Coherence**

by

Stephen T. Hamilton

Thesis submitted for the degree of Doctor of Philosophy

Department of Psychology

University of Sussex

July 2010

Declaration

I hereby declare that this thesis has not been submitted, either in the same or different form, to this or any other University for a degree.

Stephen Hamilton

21st January 2011

Acknowledgements

I would like to thank my supervisor, Professor Jane Oakhill, for all of her help and guidance over the years. I am enormously indebted to Jane for introducing me to academic research and reading comprehension as an undergraduate, and for her unfailing encouragement, enthusiasm and friendship throughout my studies. I would also like to thank Dr. Brendan Weekes, who has been a great colleague, supervisor and friend throughout my DPhil.

I would like to thank Professor Alan Garnham, Dr. Nicola Yuill, Dr. Evelyn Ferstl, Dr. Jessica Horst, Dr. I-Fan Su, Marika De Scalzi, Dr. Francesca Citron, Katherine Klingebiel, and Katie Twomey for their help, comments and advice in all things Psycholinguistics. Extra thanks to Evelyn Ferstl for introducing me to the 'CoCo paradigm' that features prominently in Chapter 5.

I would like to thank all of the staff and faculty in the Psychology department at The University of Sussex for everything that they have done for me over the years. Sussex has been a wonderful place to study, and I will always be proud to have been educated here.

Special thanks go to the Laines Lunch Gang; Drs. Sophie George, Sarah Garfinkel, Hugh Webster, Ben Charlton, and Lucy Chambers for sharing the pleasures of donut flavoured chips and wobbly quiche with me, and for their bottomless capacity for debate and banter. I would especially like to thank my fellow expats, Sophie and Sarah, for answering the phone and being amazing.

Of course, I would like to thank all of my friends in Brighton and beyond... Special thanks go to Nat Davies and Tony King for being great.

I would like to thank my parents for their love, support, and Sunday morning motivational speeches. Thanks to my brothers; Josh, Eamon, and Ross for all of the dead arms and legs, and my sisters-in-law; Rachel, Koren and Rea for being the best sisters I could have hoped for. Thanks also to my nieces and nephew; Freya, Pippa, Ayla, and Tom.

Last, I would like to thank my A-level Psychology and English Language teachers; Tim Stocker, Andrea Hulme, and Simon Fielding for introducing me to subjects that interested me and made sense to me.

This research was supported by the Economic and Social Research Council.

University of Sussex

Stephen T. Hamilton

Doctor of Philosophy

Reading Comprehension in Adults:Component Skills; False Memories; and Judgements of CoherenceSummary

The aim of this thesis was to investigate some of the processes that contribute to the effective comprehension of text in an adult population. The thesis begins with an assessment of component skills that are of theoretical relevance to reading comprehension skill.

Experiment One explored the relation between gist-based memory processes and reading comprehension skill. Weaknesses in semantic processing have been shown to contribute to comprehension difficulties both in childhood (e.g. Nation & Snowling, 1999), and adulthood (e.g. Perfetti, Yang & Schmalhoffer, 2007). Weekes, Hamilton, Oakhill & Holliday (2008) used the false memory (DRM) paradigm developed by Deese (1959); Roediger and McDermott (1995) to assess the relation between reading comprehension and memory processes in children. In the DRM, subjects memorise lists of semantically related words (e.g. *bed, rest, awake*) for later recall. During recall, it is typical to see intrusions of semantically related but non-presented items (e.g. 'sleep' is often falsely recalled following presentation of the above). Weekes et al. (2008) found that children with comprehension difficulties produced fewer such intrusions than did good comprehenders, suggesting that poor comprehenders have difficulty extracting the central theme or 'gist' from the word lists, a deficit that was attributed to weakness in semantic processing and memory. Experiment One demonstrated that this effect was not replicable in an adult population. Although there is evidence that deficits in semantic

processing contribute to reading comprehension difficulties in adulthood, these appear to be too subtle to manifest themselves in the DRM paradigm.

In Experiment Two, measures of vocabulary, word-level skills (orthography and decoding), working memory and verbal IQ were taken from a population of young adult readers. These measures were used as predictors of comprehension skill in multiple regression analyses. Moderate support for the Verbal Efficiency/Lexical Quality Hypothesis (Perfetti, 1985; 2007) was obtained, in that word-level skills and vocabulary size accounted for unique portions of variance in comprehension skill.

Experiments Three and Four explored the processes involved in on-line reading comprehension and, specifically, in a comprehension task that demanded integration. In both experiments, subjects took part in a coherence judgement task (Ferstl, Guthke & von Cramon, 2002; Ferstl, 2006) in which they had to verify whether two sentences cohered with one another or not. Four conditions that resulted from crossing coherence and cohesion (i.e. the presence of a lexical connection), were used: Coherent and cohesive (where sentences cohered, and a cohesive tie made their coherence explicit); coherent and incohesive (where sentences cohered, but coherence had to be inferred on the basis of pragmatic information rather than lexical cohesion); incoherent and cohesive (where sentences that do not cohere were erroneously linked with a cohesive tie); and incoherent and incohesive (where sentences did not cohere, and were not erroneously linked with a cohesive tie). Typically, the paradigm elicits an interaction between coherence and cohesion in reading times for the second (target) sentence: Targets in coherent and cohesive trials are read more quickly than targets in coherent and incohesive trials; and targets in incoherent and incohesive trials are read more quickly than are targets in incoherent and cohesive trials. Experiment Three replicated this interaction, and demonstrated that variance in its size was predicted by working memory capacity, with high working memory readers showing larger interaction effect sizes than low capacity readers. The interaction was interpreted as a monitoring effect that was triggered by target sentences in the atypical conditions (i.e. incoherent and cohesive; coherent and incohesive). It was proposed that high capacity readers were better able to engage in this

monitoring. Experiment Four sought to explore the semantic deficit hypothesis in relation to this effect, with the proposal that efficient semantic processes, rather than working memory capacity, contributed to variance in the size of the interaction. Performance on a semantic fluency task was found to predict unique variance in the size of the interaction effect, over and above that accounted for by working memory capacity. This finding suggests that the effect is better explained by semantic processing than by working memory capacity, and that the interaction may be better described as a semantic elaboration effect rather than a comprehension monitoring effect.

The conclusion of this thesis is that reading comprehension in adult readers relies upon efficient and accurate lexical access, comprising both lower-level processes such as accurate word recognition and decoding skill, and higher-level processes of semantic elaboration and integration.

Table of Contents

| | | |
|----------|---|-----------|
| 1 | INTRODUCTION | 1 |
| 2 | LITERATURE REVIEW | 6 |
| 2.1 | CHAPTER OVERVIEW | 6 |
| 2.2 | THE CONSTRUCTION-INTEGRATION MODEL | 7 |
| 2.2.1 | COHERENCE AND COHESION..... | 9 |
| 2.2.2 | THE MICRO- AND MACRO- STRUCTURES..... | 10 |
| 2.3 | WORD-LEVEL INFLUENCES..... | 11 |
| 2.4 | THE LEXICAL QUALITY HYPOTHESIS..... | 12 |
| 2.5 | WORKING MEMORY | 13 |
| 2.6 | LONG-TERM WORKING MEMORY AND BACKGROUND KNOWLEDGE | 17 |
| 2.7 | THE CONNECTIONIST ACCOUNT OF WORKING MEMORY AND READING SKILL..... | 21 |
| 2.8 | THE PRESENT STUDIES..... | 22 |
| 3 | FALSE MEMORIES AND READING COMPREHENSION..... | 28 |
| 3.1 | CHAPTER OVERVIEW | 28 |
| 3.2 | EXPERIMENT 1: FALSE MEMORY EFFECTS IN YOUNG ADULTS | 29 |
| 3.2.1 | INTRODUCTION | 29 |
| 3.2.2 | METHOD..... | 32 |
| 3.2.3 | RESULTS..... | 38 |
| 3.2.4 | CONCLUSIONS..... | 45 |

| | | |
|----------|---|-----------|
| 4 | COMPONENT SKILLS OF COMPREHENSION | 48 |
| 4.1 | CHAPTER OVERVIEW | 48 |
| 4.2 | EXPERIMENT 2: WORD-LEVEL ABILITY, VOCABULARY, AND WORKING MEMORY AS PREDICTORS OF READING COMPREHENSION SKILL..... | 50 |
| 4.2.1 | INTRODUCTION | 50 |
| 4.2.2 | METHOD..... | 56 |
| 4.2.3 | RESULTS..... | 58 |
| 4.2.4 | CONCLUSIONS..... | 63 |
| 5 | COMPONENT SKILLS, SEMANTIC ACCESS AND JUDGEMENTS OF COHERENCE | 66 |
| 5.1 | CHAPTER OVERVIEW | 66 |
| 5.2 | GENERAL INTRODUCTION..... | 67 |
| 5.3 | EXPERIMENT 3: COHERENCE, COHESION AND COMPONENT SKILLS | 73 |
| 5.3.1 | INTRODUCTION | 73 |
| 5.3.2 | METHOD..... | 77 |
| 5.3.3 | RESULTS..... | 80 |
| 5.3.4 | CONCLUSIONS..... | 90 |
| 5.4 | EXPERIMENT 4: COHERENCE, COHESION, COMPONENT SKILLS, SEMANTIC ACCESS AND VERBAL INTELLIGENCE | 94 |
| 5.4.1 | INTRODUCTION | 94 |
| 5.4.2 | METHOD..... | 102 |
| 5.4.3 | RESULTS..... | 107 |

| | | |
|----------|--|------------|
| 5.4.4 | CONCLUSIONS..... | 116 |
| 5.4.5 | COMBINED ANALYSES..... | 114 |
| 5.5 | GENERAL DISCUSSION..... | 121 |
| 6 | SUMMARY AND GENERAL DISCUSSION | 132 |
| 6.1 | CHAPTER OVERVIEW | 132 |
| 6.2 | SUMMARY OF FINDINGS | 132 |
| 6.3 | THEORETICAL IMPLICATIONS | 135 |
| 6.3.1 | GIST-BASED MEMORY PROCESSES | 135 |
| 6.3.2 | COMPONENT SKILLS | 137 |
| 6.3.3 | WORKING MEMORY AS A COMPONENT OF READING COMPREHENSION..... | 139 |
| 6.3.4 | COHERENCE, COHESION, WORKING MEMORY AND SEMANTIC FLUENCY | 140 |
| 6.4 | CONCLUSIONS..... | 142 |
| 6.5 | FUTURE DIRECTIONS..... | 145 |
| 6.6 | FINAL CONSIDERATIONS..... | 147 |
| | REFERENCES..... | 148 |
| | APPENDIX 1: MATERIAL USED IN CHAPTER 3 (EXPERIMENT 1) | 155 |
| | APPENDIX 2: MATERIAL USED IN CHAPTERS 3, 4, AND 5 (EXPERIMENTS 1, 2, 3, AND 4)..... | 157 |
| | APPENDIX 3: MATERIAL USED IN CHAPTER 5 (EXPERIMENTS 3 AND 4)..... | 167 |

List of Figures and Tables

| | |
|---|-----|
| Table 1. Descriptive statistics | 39 |
| Table 2. Component reading characteristics of the two comprehension groups | 40 |
| Figure 1. Recall performance of good and poor comprehenders in the semantic DRM..... | 41 |
| Figure 2. Recall performance of good and poor comprehenders in the phonological DRM..... | 41 |
| Figure 3. Recognition A' values for good and poor comprehenders in the semantic DRM..... | 43 |
| Figure 4. Recognition A' values for good and poor comprehenders in the phonological DRM..... | 44 |
| Table 3. Descriptive statistics..... | 58 |
| Table 4. Pearson's correlation coefficients..... | 58 |
| Table 5. Pearson's correlation coefficients..... | 59 |
| Table 6. Stepwise hierarchical multiple regression analysis with ND comprehension as the dependent variable..... | 60 |
| Table 7. Stepwise hierarchical multiple regression analysis with ND comprehension as the dependent variable..... | 62 |
| Table 8. Example materials from the four conditions of the coherence and cohesion paradigm | 71 |
| Table 9. Descriptive Statistics..... | 80 |
| Table 10. Accuracy and reading times for all participants in all conditions of the coherence and cohesion paradigm..... | 81 |
| Figure 5. Mean percentage error rates for all conditions of the coherence and cohesion paradigm..... | 81 |
| Figure 6. Mean standardised residuals of target sentence reading times in the coherence and cohesion paradigm..... | 84 |
| Table 11. Component reading skills of working memory groups..... | 86 |
| Figure 7. Mean residuals of target sentence reading times for the high working memory group (n=28) and the low working memory group (n=24). Error bars represent one standard error above and below the mean..... | 87 |
| Table 12. Pearson's's correlation coefficients between component variables and interaction value | 89 |
| Table 13. Stepwise hierarchical regression analysis with absolute interaction value as the dependent variable..... | 90 |
| Table 14. Descriptive statistics..... | 107 |
| Table 15. Accuracy and reading times for all participants in all conditions..... | 108 |
| Figure 8. Mean percentage error rates for all conditions (n=63). Error bars indicate one standard error above and below the mean..... | 108 |
| Figure 9. Mean standardised residuals of target sentence reading times (n=63). Error bars indicate one standard error above and below the mean..... | 110 |

| | |
|--|-----|
| Table 16. Pearson's correlation coefficients..... | 113 |
| Table 17. Stepwise hierarchical regression with absolute interaction value as the dependent variable..... | 115 |
| Table 18. Descriptive statistics..... | 118 |
| Table 19. Accuracy and reading times for all participants in all conditions..... | 119 |
| Figure 10. Mean percentage error rates for all conditions..... | 119 |
| Figure 11. Mean standardised residuals of target sentence reading times..... | 121 |
| Table 20. Component reading skills of the two working memory groups..... | 122 |
| Figure 12. Mean residuals of target sentence reading times for the two working memory groups..... | 123 |
| Table 21. Pearson's correlation coefficients between the component variables and the interaction values..... | 124 |
| Table 22. Regression analysis with absolute interaction value as the dependent variable..... | 125 |
| Table 23. Span characteristics of the low and high working memory groups in Experiments 3 and 4..... | 126 |

1 Introduction

Reading is a complex intellectual skill that involves transforming written symbols into meaningful mental representations. Reading skill is a prerequisite for success in our society, but our motivations to read are driven by more than our desire for financial autonomy and professional accomplishment. We read to garner information and develop our knowledge of the world, to enlighten ourselves to the opinions and perspectives of others, to engage emotions that we may not wish to experience in our daily lives, to acquire vocabulary and expertise, and to entertain ourselves. Reading comprehension is the process by which we come to understand a text, and allows us to fulfil these motivations. It is why we read, it is why we teach others to read, and it is why we value reading so highly.

Comprehension occurs as a reader builds a mental representation of a text's message. This extracted mental representation is referred to as a *situation* or *discourse* model of a text (Johnson-Laird, 1983; Kintsch, 1988; Graesser, Singer & Trabasso, 1994). The discourse model comprises information that is stated explicitly in the text and information that is *inferred* by the reader on the basis of knowledge. For example, the statement, "James reached into his bag and pulled out the flask. The orange juice tasted delicious," requires at least two inferences in order to be comprehensible: That the flask contained orange juice, and that James opened the flask and drank its contents.

For skilled readers, reading is usually experienced as automatic and effortless. The apparent simplicity of the reading process belies an extremely complex cognitive ability

that relies on a multitude of skills for its execution. To achieve comprehension, one must decode the symbols given by a written orthography, map those symbols onto representations in the mental lexicon, derive the meanings of individual words, and integrate those meanings with one another on the basis of syntactic information and background knowledge. Both word meanings and world knowledge are integrated very rapidly, within around 400ms of encountering a word (Hagoort, Hald, Bastiaansen, & Petersson, 2004). That the reading process operates so quickly and with such apparent ease is testament to our ability to learn. Unlike speaking and listening comprehension, reading comprehension is not something to which our brain has evolved, and it often requires deliberate instruction to be adequately accomplished. It should come as no surprise, therefore, that reading comprehension skill varies enormously between individuals (Kirby, 2007), even among well-educated college students (Bell & Perfetti, 1994).

The aim of this thesis was to investigate some of the processes that contribute to the effective comprehension of text in an adult population. For this purpose, participants were recruited from undergraduate and postgraduate student populations, as well as members of the general population who had attained a university-level of education.

Variance in reading comprehension skill can be attributed to a number of factors. Children with comprehension problems are known to have weaknesses in semantic processing (e.g. Nation & Snowling, 1999) and have difficulties integrating general knowledge and text information (Yuill & Oakhill, 1991). There is evidence to suggest that

these problems may be related to difficulties in extracting the gist meaning from text (Weekes, Hamilton, Oakhill & Holliday, 2008).

When reading, an individual must store pragmatic, semantic and syntactic information given by the text, and use it to parse and integrate subsequent text into a coherent mental representation. These processes are believed to depend on working memory; a limited capacity temporary storage space used for the storage and manipulation of incoming information (Baddeley, 2003). Since working memory capacity is known to be limited (e.g. Daneman & Carpenter, 1980), it is thought that variance in comprehension skill can be attributed to variance in this capacity (Just & Carpenter, 1992). Alternative theories of working memory (e.g. Ericsson & Kintsch, 1995; MacDonald & Christiansen, 2002) posit that the relationship between working memory and reading comprehension is based on linguistic experience, which determines the efficiency and aptitude with which the cognitive system can deal with linguistic information.

It is known that adults with reading comprehension difficulties have slower and inefficient word-level skills (Bell & Perfetti, 1994), and it is proposed that this laboured process of lexical access affects reading comprehension by monopolising working memory resources that would otherwise be allocated to higher-order comprehension processes such as inference generation and sentence integration (Perfetti & Hart, 2002; Perfetti, 2007).

When readers encounter a gap in coherence they must reflect upon their knowledge of the concepts that are engendered by the text, and engage in inferencing processes that go beyond what is stated explicitly. This process of elaboration and

integration, dependent upon knowledge, allows them to maintain and develop a coherent mental representation of the text's message (McNamara, 2001). If the reader is unable to access this knowledge and integrate it with their current mental representation, then comprehension will suffer as a consequence.

In preparation for the empirical research, Chapter 2 begins with a presentation of literature that is relevant to reading comprehension skill in adults, with an emphasis on the Construction-Integration model of text comprehension (Kintsch, 1988; 1998), the Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007), and three models of working memory (Just & Carpenter, 1992; Eriksson & Kintsch, 1995; MacDonald & Christiansen, 2002). In Chapter 3, gist-based memory processes and their relation to reading comprehension are discussed. Experiment 1 uses the Deese/Roediger McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995) to explore the relation between gist-based memory processes and reading comprehension, with the contention that false recollection of semantically related but non-presented items will be related to performance on a standardised measure of comprehension skill. Contrary to the findings of Weekes et al. (2008), false recollection is not related to comprehension skill in the present sample of adult readers. The results are discussed in relation to background knowledge and the processes of construction and integration (Kintsch, 1988; 1998). In Experiment 2, component skills of comprehension (word-level ability; vocabulary; working memory capacity and verbal intelligence) are measured in a sample of university students, and their ability to predict variance in comprehension skill is ascertained through multiple regression analyses. Moderate support for the Lexical Quality Hypothesis (Perfetti & Hart,

2002; Perfetti, 2007) is obtained, in that word-level skills and vocabulary size account for unique portions of variance in comprehension skill. In Chapter 5, a discussion of on-line comprehension skill and coherence building is presented, and an on-line coherence judgement task (Ferstl, Guthke & von Cramon, 2002; Ferstl, 2006) is introduced. This task is used to investigate component skills that may be related to integration processes that operate across sentence boundaries. In Experiment 3, comprehension skill, vocabulary, working memory and word-level ability are used as predictors of variance in performance on the coherence judgement task. This experiment shows that working memory capacity is highly related to variance in standardised reading times in the paradigm, a finding that is explained with reference to the capacity constrained model of comprehension (Just & Carpenter, 1992). In Experiment 4, additional measures of semantic access and verbal intelligence are taken to further explicate the process of integration. Multiple regression analyses demonstrate semantic fluency to be highly related to integration processes in the coherence judgement paradigm. This finding is explained with reference to models of working memory that emphasise the importance of linguistic experience in relation to comprehension skill.

The thesis closes by discussing of the importance of word-level ability, vocabulary, executive functions and semantic processing to reading comprehension in adults, and suggests that future research should aim to elucidate the contribution of language experience to performance on these tasks and to variance in adult comprehension skill more generally.

2 Literature Review

2.1 Chapter Overview

Reading is a complex intellectual skill whose execution relies upon multiple component processes. Large individual differences in reading skill are seen even among university students (Bell & Perfetti, 1994; Braze, Tabor, Shankweiler, & Mencl, 2007). Reading comprehension involves language-specific processes as well as domain-general cognitive abilities such as sensation, perception, attention, memory, and reasoning. Variation in any one of these abilities potentially underlies individual differences in comprehension performance (Long, Johns, & Morris, 2006).

When reading, comprehension occurs as the reader builds a mental representation of the text's message. In the current literature, this extracted mental representation is referred to as the *situation or discourse model* of a text (Johnson-Laird, 1983; Kintsch, 1988; Graesser, Singer, & Trabasso, 1994). The situation model comprises information that is mentioned explicitly in the text and information that is inferred by the reader on the basis of world knowledge.

This chapter begins with a description of the Construction-Integration (CI) model, with an emphasis on the processes that may be affected by individual differences. The Construction-Integration model (Kintsch, 1988; 1998), gives a detailed and modular account of the comprehension process. In doing so, it provides opportunities for the identification of where the process may be constrained by individual differences.

The CI model emphasises the importance of knowledge to the reading process. Knowledge is drawn upon in to maintain coherence in the face of coherence gaps, and is used to produce a global representation of a text. This process may be mediated by working memory capacity/function. In this chapter, three models of working memory are presented in some depth, and their ability to explain individual differences in reading comprehension skill is evaluated.

The CI model proposes that when words are encountered, readers map lexical representations for those words onto nodes for concepts or propositions which are then used to construct mental representations of texts. The CI model does not describe how lexical access may differ between individuals and it does not describe how low-level processes such as decoding may impact on the reading process. The Lexical Quality Hypothesis (Perfetti & Hart, 2002, Perfetti, 2007) specifies that differences in reading skill are attributable to differences in word-level skills, and is described in some depth.

This chapter concludes by introducing the experiments that were designed to investigate reading comprehension in adult readers in light of the proposals and predictions of the presented literature.

2.2 The Construction-Integration Model (Kintsch, 1988; 1998)

According to the CI model, texts are represented at three levels of increasing sophistication: The *surface structure*, the *text-base* and the *situation* or *discourse model*. The surface structure encodes the verbatim form of the text, providing orthographic and phonological representations for words, and syntactic relations between them. The

surface structure is then used to create the text-base via a process of *construction*.

Construction refers to the activation of information, both in the text and in an associative network of the reader's knowledge. The term *knowledge net* is used to describe this associative network, made up of nodes for concepts and propositions that are linked to one-another via associative links (Kintsch, 1998). Construction is based on associative activation, with items in working memory activating propositions in the knowledge net via bottom-up, automatic, retrieval-based processes akin to priming (Kintsch, 1998). This construction is assumed to be automatic, and is not constrained by top-down processes such that *all* knowledge related to a word is activated. It is through a process of *integration* that this activation is directed toward relevant information (Kintsch, 1988).

Working memory is known to be limited (e.g. Daneman & Carpenter, 1980), but because nodes in the knowledge net are connected through *retrieval structures* to other nodes in the net, very complex ideas can be generated and represented even though only a few nodes are activated at any one point (Kintsch, 1998). The activation of nodes in working memory determines how they will activate these retrieval structures and their associated concepts. Activated concepts will activate other concepts via spreading activation. Integration refers to the spreading of activation across the network until it stabilises, resulting in pronounced activation of information that is pertinent to the ideas conveyed by the text, and reduced activation of concepts that are irrelevant to the text representation (Kintsch, 1988).

The CI model assumes that the fundamental unit of processing in the text-base representation is the proposition. A proposition consists of a predicate and its arguments

(Kintsch, 1998). For example, the sentence, “The chef baked a cake,” would be propositionally represented as, BAKE (CHEF, CAKE). Links between propositions in the text-base representation are given by *argument overlap*, which provide explicit connections between ideas in the text (McNamara & Kintsch, 1996). When arguments between propositions do not overlap a gap in coherence is created, requiring the reader to produce a bridging inference to maintain textual coherence. For example, the sentence above may be followed by, “Angered by the sight of the cherry on top, the chef pulverised it with his bare fists, covering the walls with chocolate, cream and strawberry jam.” The only argument that overlaps between the two sentences is *chef*. To maintain coherence, an inference needs to be generated on the basis of knowledge (e.g. that *it* is the *cake* of the preceding sentence). When readers generate bridging and knowledge-based inferences that go beyond the text and incorporate them into their mental representation, the representation is said to be at the level of the *situation model*.

The situation model includes information about the characters, settings, events and actions that are either stated explicitly or are implied by the text (Graesser et al., 1994). The situation model and the text base are not different or separate mental representations; they are different dimensions of episodic memory for the text (Kintsch, 1988).

2.2.1 Coherence and Cohesion

A mental representation that is coherent is one in which text-derived propositions are successfully connected to one-another via local and global inferences. As stated, these

inferences are made on the basis of knowledge. Links to knowledge are assumed to be made throughout the reading process, but it is also proposed that breaks in the discourse induce the reader to reflect upon and to activate more background knowledge more deeply in order to maintain coherence (Kintsch, 1998). McNamara and colleagues have obtained evidence to show that text cohesion and prior knowledge *interact* (McNamara, Kintsch, Songer, & Kintsch, 1996; McNamara, 2001). That is, low knowledge readers benefit from greater cohesion (argument overlap, the presence of cohesive ties that serve to link adjacent clauses, etc.) because they lack the necessary knowledge to generate bridging inferences autonomously (McNamara & Magliano, 2009). High-knowledge readers, on the other hand, benefit most from texts that are of low cohesion: Coherence gaps that are produced by the omission of cohesive devices force them to reflect upon prior knowledge to a greater degree, thereby producing inferences that contribute to the development of an elaborated text-base and situation model (McNamara, 2001).

2.2.2 The Micro- and Macro- Structures

Texts are not only represented at the level of the surface structure, the text-base and discourse level, they are also represented at a local level (the *microstructure*), and a global level (the *macrostructure*). The microstructure consists of the propositions within the text and their interrelations (including local or bridging inferences). The macrostructure is a hierarchical organisation of the propositions in the microstructure (Kintsch, 1998). This hierarchical representation imparts the global organisation of the text; the main ideas, themes, or *gist* of the text. A number of factors are crucial to the

formation of macrostructures, including general cultural knowledge (e.g. knowledge of cakes), the participant categories (e.g. chefs), and the type of text (e.g. expository vs. narrative) (Kintsch, 1998). It is known that children with comprehension difficulties have problems integrating general knowledge with textual information to fill in missing details (e.g. Baker & Stein, 1981), as well as integrating ideas in a text in order to answer questions about main ideas and themes (Yuill & Oakhill, 1991). It is thought that these difficulties may, in part, arise because of semantic processing deficits (e.g. Nation & Snowling, 1999), and may be due to deficits in gist-based memory processes (Weekes, Hamilton, Oakhill, & Holiday, 2008).

2.3 Word-level influences

Readers are unlikely to make inferences to integrate or elaborate ideas in a text if they fail to construct an accurate propositional, or text base representation of it. As stated above, the construction of the text base depends upon access to nodes for concepts and propositions of a text via knowledge. This access is granted by an accurate representation of the surface structure of the text. The construction of the surface structure relies upon a foundational processing factor: The accurate identification of words. For successful comprehension, readers need to access the meaning of a given word. Such identification not only requires accuracy (knowing word meanings), but also fluency (speed of lexical access), and richness (semantic network connections) (Beck, Perfetti, & McKeown, 1982).

2.4 The Lexical Quality Hypothesis

Adults with reading comprehension problems have been shown to have slower and less efficient word identification skills, as well as difficulties processing low frequency syntactic structures (Bell & Perfetti, 1994). The so-called Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) holds that the higher-order process of reading comprehension relies upon efficient and accurate lexical access. That is, reading comprehension problems result from impoverished lexical representations that constrain higher-order comprehension processes.

A high-quality lexical code should have orthographic, phonological and semantic specificity, allowing a word to be processed and understood coherently when it is encountered in both visual and auditory modalities (Perfetti, 1985). In this model, coherence refers to the idea that lexical features codes are accessed in synchrony, resulting in a unitary word perception event (Perfetti & Hart, 2002). Perfetti and colleagues argue that multiple encounters with a word will produce a common 'core representation' that provides a nexus between orthographic, phonological, and semantic information of a word.

According to the lexical quality hypothesis, inferior lexical codes affect reading comprehension in three ways. First, the comprehension processes required to construct an accurate text base, such as syntactic analysis, will be negatively affected by low quality lexical codes, as semantic specificity cannot be established (e.g. morphological features that indicate tense or plurality may not be properly encoded). Second, the comprehension processes needed to construct discourse representations at the situation level (i.e.

knowledge-based inferences) will be impeded by lexical representations that are impoverished at the semantic level. That is, retrieval-based knowledge activation (i.e. *construction*) that is actuated by poorly understood words will produce an inadequate collection of concepts that can be used to generate knowledge-based inferences. Last, resources that would normally be dedicated to higher-order comprehension processes (i.e. processes such as inference generation and sentence integration) will be consumed by lexical processing. This final proposal reflects the role working memory plays in reading comprehension: allocating cognitive resources to the task.

2.5 Working Memory

When reading a text, one must store pragmatic, semantic and syntactic information given by the text and use it to disambiguate, parse, and integrate subsequent text into a coherent mental representation (Daneman & Carpenter, 1980). These processes are believed, in part, to be dependent upon working memory; a limited capacity temporary space used for the storage and manipulation of incoming information (Baddeley, 2003). Working memory is the theoretical construct that is used to refer to the system or mechanism underlying the maintenance of relevant information during the performance of a cognitive task (Baddeley & Hitch, 1974). Baddeley and Hitch (1974) proposed that working memory can be divided into three subsystems; the phonological loop (concerned with verbal and acoustic information); the visuospatial sketchpad (concerned with the storage of visual information); and third, an attentionally-limited control system - the central executive (Baddeley, 2003).

Traditional measures of short term memory such as digit span (how many single digits one can keep in mind at any one time) and word span (how many words one can keep in mind at any one time) are only weakly correlated with reading ability (Daneman & Carpenter, 1980). One explanation is that these tasks only tap short-term *storage* processes of the phonological loop (Whitney, Arnett, Driver, & Budd, 2001), and since working memory is assumed to have processing as well as storage functions, individual differences in simple span measures cannot account for individual differences in cognitively demanding tasks that are characterised by storage and processing requirements.

A task that taps both processing and storage components of working memory was developed by Daneman and Carpenter in 1980. The reading span task requires participants to read aloud or listen to groups of sentences, half of which make sense and half of which do not. Subjects have to state whether or not each sentence makes sense (the processing component) and to store the last word from each of the sentences; producing them in the order in which they were presented at recall. The number of sentences in a set is incrementally increased from trial to trial, increasing the load on working memory. The participant's 'reading span' is taken as the maximum number of sentences they are able to judge correctly, while maintaining perfect recall of the sentence-final words. The reading span task requires subjects to maintain to-be-remembered (TBR) items in the phonological loop; to process incoming information for judgements of coherence; and to inhibit irrelevant information at recall (e.g. items that had been presented in earlier lists). The reading span task is thought to give a measure of

processes attributable to both the phonological loop and to the processing and inhibition functions of the central executive (Whitney et al., 2001).

There is a broad range of ability in the task, even among well-educated college students; recall performances ranging from 2 to 5 items were recorded by Daneman and Carpenter (1980) in their cohort of undergraduate subjects. The task has now been developed to include up to 7 TBR items (e.g. Unsworth, Redick, Heitz, Broadway, & Engle, 2009). As working memory capacity is limited, individual differences in comprehension skill are attributed by some to individual differences in capacity (e.g. Just & Carpenter, 1992). Indeed, Daneman and Carpenter (1980) found that performance on the reading span task was significantly correlated with a number of comprehension measures, including Verbal Scholastic Aptitude Test scores ($r = .59$); fact retrieval ($r = .72$) and accuracy in pronominal reference resolution ($r = .90$).

Many theories propose that differences in working memory capacity depend upon differences in the amount of *activation* that is available its storage and processing components (e.g. Just & Carpenter, 1992; Lovett, Reder, & Lebiere, 1999). Implicit in this contention is that the amount of activation required for the processing of a particular text is dictated by the characteristics of the text itself. For example, if the reader is expecting to encounter a noun phrase after reading a transitive verb (e.g. “John ate...”), they may also be calculating other, syntactic, pragmatic or semantic features of that sentence. If the number of these calculations is large (e.g. when a sentence is syntactically complex, or there are a large number of propositions that need to be processed and maintained) then this activation will be distributed across the processes, resulting in scaled-down processing

of incoming information. That is, when task demands exceed the constraints of available resources, both storage and computational functions of working memory are degraded. This 'capacity-constrained' theory of comprehension (Just & Carpenter, 1992) holds that reading skill differences are, in part, due to individual differences in the amount of activation that is available to the language processing system. Under this proposal, each representational element (a word, phrase, proposition, grammatical structure, thematic structure, object in the external world, etc.) has an associated activation level. If this element's activation level is above some minimum threshold it can be operated on by processes in working memory (Just & Carpenter, 1992). Thus, if the total amount of activation available to the system is less than that required to perform the task then some elements will be displaced through lack of activation (Just & Carpenter, 1992).

Although working memory is conceived by activation theories as a unitary construct whose functioning is attenuated by the availability of resources, the cost of representing and maintaining elements of a text will, presumably, vary as a function of a reader's experience with those elements. Like Just and carpenter (1992), Lovett, Reder and Lebiere (1999) propose that working memory is a cognitive resource that, (a) can be allocated to the maintenance and processing of information; (b) is inherently limited; and (c) differs in supply across individuals. They propose that working memory serves to activate sources or 'nodes' in long term memory, both on the basis of its prior use, and its relevance to the current goal. Activation spreads to each available node in proportion to the strength of association between that node and the current task. Importantly, nodes are proposed to have baseline levels of activation that are dependent upon past use,

meaning that nodes with a high resting-level of activation will need little activation to be instantiated into a current task or goal.

2.6 Long -Term Working Memory and Background Knowledge

A unitary and capacity-constrained concept of working memory is thought by many to be too restrictive to explain the complex and cognitive demands of reading processes (e.g. Spilich, Vesonder, Chiesi, & Voss, 1979). Some have proposed that working memory comprises a limited capacity focus of attention that works in unison with aspects of long term memory that are accessed via the activation of retrieval cues. Ericsson and Kintsch (1995) propose a model of working memory that consists of the standard, limited capacity mechanism that they call short-term working memory, and a mechanism based on skilled storage and retrieval from long term memory that they call long-term working memory. These authors emphasise the role of practice and skill in reading processes, and argue that individuals draw on acquired knowledge and on systems of 'retrieval structures' that encode information into long term memory in a retrievable form. The structures are activated by incoming information, and act to retrieve pertinent information from LTM into a temporary storage space where it can be used to execute the task at hand (Ericsson & Kintsch, 1995).

Readers who have background knowledge relevant to the domain of a given text remember that text better and understand it more clearly than do readers who lack relevant background knowledge (Long, Prat, Johns, Morris, & Jonathan, 2008). Bransford and Johnson (1972) presented participants with text made up of seemingly random

phrases (e.g. “First, you sort them into piles based on their makeup”... “If you have to go somewhere else due to lack of facilities, that is the next step.”). Recall for the text improved dramatically when subjects were given contextual cues that denoted the passage’s meaning (i.e. the title, “How to Wash Clothes”). The inclusion of the retrieval cue granted access to the reader’s knowledge of clothes and laundry, allowing the passage to be integrated into the reader’s knowledge base via long term working memory. This access to LTM results in a more coherent and elaborated discourse representation which is then drawn upon for accurate recollection after the passage has been read.

When reading texts specific to a particular domain (e.g. Star-Trek stories (Long and Prat, 2002) and Baseball stories (Spilich, Vesonder, Chiesi and Voss, 1979)), those who are knowledgeable of that domain (e.g. ‘Trekkies’) recall more information from the text than those who are less knowledgeable. It is proposed that prior knowledge helps readers to establish conceptual relations among propositions given by a text, helping readers to make knowledge-based inferences (Long & Prat, 2002). For example, if a reader has a rich representation of what it means to be a ‘clinton’ (a mythical star trek species known to be stoical), then the propositional information given by the sentence “*Worf is a Clinton,*” allows one to draw-out a richer, discourse-level (i.e. inferential) representation for the subsequent sentence, “*Lieutenant Worf remained silent*” (Long & Prat, 2002). Thus, high-knowledge readers are able to construct discourse models in which text ideas are integrated both with each other, and with a large network of prior knowledge (Long and Prat, 2002; Long, Wilson Hurley and Prat, 2006). These models can then be reflected upon to support recall, problem solving, generalisation and, as stated, knowledge based

inferences. In contrast, when reading text without access to an embellished knowledge base (relying upon short-term working memory alone), readers are only able to construct sentence-level representations, precluding the development of a higher-order discourse representation of the text.

Long and her colleagues (Long, Johns, Morris, & Jonathan, 2008) investigated how background knowledge is related to text memory, using remember/know judgements of text ideas as their dependent measure. Recollection estimates to old items were predicted by readers' background knowledge, but *not* by word-decoding skill and working memory capacity. Furthermore, false positive responses to new items that were related to the text inferentially were *diminished* as a function of verbal ability, working memory capacity and reasoning, but *increased* as a function of background knowledge (Long et al., 2008). Thus, background knowledge supports recollection of texts and drives knowledge-based inferences, presumably through a process whereby text information is integrated with background knowledge. Experiment 1 in the present thesis explored this idea with the use of the DRM (Deese, 1959; Roediger & McDermott, 1995) 'false-memory' paradigm. Weekes, Hamilton, Oakhill, & Holiday (2008) found that children with comprehension difficulties produced fewer extra-list intrusions during recall of semantically related word lists than did good comprehenders. This suggests that poor comprehenders have difficulties linking incoming information with background knowledge.

The findings presented above suggest that poor comprehenders' problems may be caused by deficient or poorly constructed text bases that are devoid of the propositional representations needed to make inferences (Long & Prat, 2002). However, in children at

least, there is evidence that this may not be quite so straightforward. Cain, Oakhill, Barnes and Bryant (2001) taught two groups of good and poor comprehenders a knowledge base about a fictional world called 'Gan' (e.g. "In the world of Gan, turtles have ice skates attached to their feet," "the bears on Gan have bright blue fur."). The children were then read a story about 'Gan' and asked comprehension questions, some of which tapped their ability to produce coherence inferences. For example, in one story the protagonists had trouble walking down an icy path, when one character states, "I wish I was a turtle!" - A statement can be understood only if the information from the knowledge base was used for its interpretation. The child was asked, "what did Dack wish?" - A question that could only be answered properly if the child had made an inference for coherence. Other questions tapped the children's ability to make elaborative inferences. For example, in one story the protagonists pull coats made of bear fur out of their bag. The child was asked "what did Dack and Dane pull out of their bags?" If the child had made an elaborative inference then they would have stated that the coats were blue. Although this inference is not necessary for comprehension, it would create a richer representation of the text. Poor comprehenders had difficulty making both types of inference, even when differences in memory for the knowledge base and differences in memory for the text itself were accounted for. Thus, children with reading comprehension difficulties show problems forming a coherent discourse representation even when they have adequate knowledge with which to do so. The authors acknowledge that the study did not control for *ease of access* to knowledge, and it may be that poor comprehenders were restricted by the accessibility of information in the taught knowledge base. Sure enough, a common

source of inference failure for the less skilled comprehenders was a failure to retrieve the correct textual premise (e.g. a failure to recall that Dack wished he was a turtle). Conversely, inference failures in the skilled comprehenders occurred at a different level; often they recalled the correct premise and the correct information from the knowledge base, but failed to integrate the two. The authors conclude that the likely source of inference failure in the less-skilled comprehenders was an inability to select the information relevant to make the inference. Potentially, selecting the correct information for an inference is a process that is sub-served by working memory or executive functioning.

2.7 The Connectionist Account of Working Memory and Reading Skill

Alternatively, individual differences in working memory and comprehension skill have been attributed to individual differences in skill and knowledge rather than differences in capacity per se. That is, it is the *functioning* of working memory that differs between good and poor comprehenders. The connectionist-based account of variation in working memory (MacDonald & Christiansen, 2002) proposes that working memory capacity varies as a consequence of its functional architecture (i.e. the number of processing units, how activation passes through weights etc.) and that this is modulated by the network's experience (i.e. how often it has processed similar input in the past) (Long, Johns, & Morris, 2006). Consequently, differences in reading skill do not stem from variations in working memory capacity, instead they emerge from biological factors and

experience with language. MacDonald and Christiansen (2002) argue that the distinction that is drawn between language processing tasks and linguistic working memory tasks is an artificial one, and that all of these tasks are simply different measures of language processing skill. Further, differences between individuals can not only be attributed to variations in exposure to language and reading, but to biological differences that affect processing accuracy (e.g. differences in the precision of phonological representations) (MacDonald & Christiansen, 2002).

The network's ability to process incoming information varies as a function of the input (e.g. whether the material is complex or simple), the properties of the network (how activation is passed through weights etc.), and the interaction of these properties – how much the network has experienced similar input before (MacDonald & Christiansen, 2002). Although similar to the activation theories presented above, the key distinction is that the connectionist account proposes that there is no working memory storage and manipulation space per se. Rather, working memory resides within the network itself; it is not a separate entity that can vary independently of the architecture and experience that governs the network's processing efficiency (MacDonald & Christiansen, 2002).

2.8 The Present Studies

The CI model does not describe how comprehension skill may differ between individuals, but aspects of the model allow for some predictions to be made. First, the fact that the limited nature of working memory constrains the number of nodes (i.e. propositions) that can be activated at any one time suggests working memory capacity

should act to constrain comprehension processes. This forms the central thesis of the *capacity-constrained* model comprehension of Just and Carpenter (1992). However, Kintsch (1998) maintains that nodes are linked to long term memory via retrieval structures, meaning that very complex ideas can still be represented in the face of a limited capacity working memory. If the activation of these nodes is to be efficient, the nodes themselves need to be well specified and their links to long term memory (via retrieval structures) need to be activated effortlessly. Further, retrieval structures themselves may be poorly represented and further, nodes for concepts in the knowledge net may be poorly networked to their associates. If so, then items in the text will not activate relevant concepts in the knowledge net, and concepts in the knowledge net will not be properly integrated with one another.

Semantic and gist-based memory processes were investigated in Experiment 1. The false memory (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995) was used to investigate the relationship between gist-based memory processes and text comprehension. In the DRM, subjects are presented with lists of semantically related items that they are asked to memorise for later recall (e.g. *Bed, Rest, Tired, duvet; Nurse, Sick, Syringe, Hospital*). At recall, subjects will often produce items that weren't presented, but were highly related to study list items (e.g. *sleep; doctor*). Weekes et al. (2008) found that false recollection in the paradigm was related to comprehension skill, with skilled comprehenders exhibiting higher rates of false recollection than poor comprehenders. Experiment 1 attempted to replicate this effect in an adult population, the hypothesis being that less skilled comprehenders would have difficulty extracting the *gist* of the

semantically related word lists, and that this is related to difficulties creating a macro-structure during text comprehension.

Although three rather different models of working memory have been presented in this chapter, the aim of this thesis was not to resolve the debate between capacity and connectionist views of working memory, but to investigate the relation between working memory and reading comprehension in adults. The sentence span task (Daneman & Carpenter, 1980) was used to assess working memory in Experiments 2, 3 and 4. Experiment 2 investigated the relation between working memory, word-level ability (decoding & vocabulary) and performance on a standardised measure of reading comprehension skill (the Nelson Denny Reading Test: Brown, Fishco, & Hanna, 1993). The capacity constrained model (Just and Carpenter, 1992) of comprehension skill would predict a direct relationship between performances on the reading span task and the comprehension component of the Nelson Denny reading test. The distributed learning model of MacDonald and Christiansen (2002) states that working memory is actually a proxy for language ability, and that working memory 'capacity' is constrained by one's proficiency in language. Therefore, this model would predict a relation between working memory and reading comprehension that is mediated by language ability (such as vocabulary and decoding skill). Last, the long term working memory model of Ericsson and Kintsch (1995), which proposes that working memory capacity is constrained by access to world knowledge, would predict a relationship between sentence span performance and comprehension that is mediated by word knowledge (i.e. vocabulary).

Experiments 3 and 4 explored the relation between working memory, decoding skill and vocabulary and performance on an on-line reading task. In both experiments, subjects took part in a coherence judgement task (Ferstl, Guthke & von Cramon, 2002; Ferstl, 2006) in which they had to verify whether two sentences cohered with one another or not. Four conditions that resulted from crossing coherence and cohesion (i.e. the presence of a lexical connection), were used: Coherent and cohesive (where sentences cohered, and a cohesive tie made their coherence explicit); coherent and incohesive (where sentences cohered, but coherence had to be inferred on the basis of pragmatic information rather than lexical cohesion); incoherent and cohesive (where sentences that do not cohere were erroneously linked with a cohesive tie); and incoherent and incohesive (where sentences did not cohere, and were not erroneously linked with a cohesive tie). Typically, the paradigm elicits an interaction between coherence and cohesion in reading times for the second (target) sentence: Targets in coherent and cohesive trials are read more quickly than targets in coherent and incohesive trials; and targets in incoherent and incohesive trials are read more quickly than are targets in incoherent and cohesive trials. The aim of Experiment 3 was to assess whether this interaction was related to working memory performance, with the proposal that the ability to establish the relation between the target and context sentences would depend on working memory capacity. Another aim was to shed light on the Lexical Quality hypothesis of reading comprehension skill (Perfetti & Hart, 2002). That is, if laboured decoding monopolises working memory resources that would otherwise be devoted to higher-order comprehension processes, such as integration, then variance in

performances and reading times on the coherence and cohesion paradigm should be, at least partially, predicted by variance in word-level skills such as decoding ability and vocabulary.

Experiment 4 sought to explore the role of semantic processing in relation to this interaction effect. The findings of McNamara and colleagues (McNamara, Kintsch, Songer, & Kintsch, 1996; McNamara, 2001) suggest that gaps in coherence are filled with access to knowledge, and the long-term working memory model of Ericsson and Kintsch (1995) predicts that access to world knowledge determines the relation between working memory and reading comprehension processes. Since the two-sentence stories used in the coherence and cohesion task referred to situations in the real world (situations that all of our participants would presumably have had experiences with) it was not possible to look at the specific relation between knowledge and performance. However, individuals may differ in their ability to *access* knowledge, such that poorer readers may show deficits in semantic access. Experiment 4 introduced a number of measures of semantic processing that were then used as predictors of the magnitude of the interaction between coherence and cohesion. The long term working memory model would predict that variance in the magnitude of this interaction would be predicted by semantic access, the capacity constrained model of working memory and comprehension would suggest that the magnitude of the interaction would only be predicted by performances on the sentence span task, and last, the lexical quality hypothesis (Perfetti & Hart, 2002) and the distributed learning model (MacDonald & Christiansen, 2002) would both predict that the

interaction effect would be at least partly predicted by variance in word-level skills such as decoding ability.

3 False Memories and Reading Comprehension

3.1 Chapter Overview

In the CI model (Kintsch, 1988; 1998) linguistic input serves to activate related concepts and propositions by a process of construction, with items from the text acting as retrieval structures for this activation. Activation extends across concepts and propositions through a process of spreading activation through a knowledge net that is unconstrained by top-down processes and is, therefore, automatic. It is known that children with reading comprehension problems have difficulties representing abstract semantic relations, and there is evidence to suggest that these children have deficits in gist-based memory processes that may be important for the construction of global representations of texts (Weekes, Hamilton, Oakhill & Holiday, 2008). The present study sought to explore this phenomenon in adult readers with the use of the DRM (Deese, 1959; Roediger & McDermott, 1995) false memory paradigm, with the hypothesis that, like children (Weekes et al., 2008), adults with reading comprehension weaknesses would show reduced rates of false recall and recognition. Contrary to expectations, no differences in false recollection between two groups of readers divided on the basis of comprehension skill were observed. Discussion of this null result centres on developmental differences and sampling issues, and concludes with the proposal that adults with comprehension weaknesses may have difficulties with the process of integration rather than construction.

3.2 Experiment 1: False Memory Effects in Young Adults

3.2.1 Introduction

False memory illusions - the phenomena of remembering events that never happened or of remembering events differently from the way that they happened – have been studied for some time. James Deese (1959) tested memory for word lists in single-trial, free recall experiments using lists containing semantic associates (e.g. *bed, rest, tired, awake*). He found that extra-list intrusions occurred frequently at recall in these experiments (e.g. *sleep* was often recalled). Further, the likelihood that an extra-list intrusion would occur increased as a function of the mean association strength of that word to the words in the study list¹. The paradigm was developed further by Roediger and McDermott (1995), and is now widely referred to as the Deese (1959); Roediger and McDermott (1995) (DRM) paradigm.

The activation-monitoring theory (Roediger & McDermott, 1999) attributes false-memory intrusions as resulting from the activation of associated words *in combination* with a failure to monitor the source of this activation. This leads participants to the erroneous conclusion that a ‘critical’, non-presented but highly related word was actually presented during the study phase of the task. An alternative account of false memory intrusions is provided by Brainerd and Reyna’s (2005) Fuzzy Trace Theory. Fuzzy trace theory is a dual process theory which holds that individuals process the surface form of

¹ Mean association strength was ascertained through free association experiments, and represents the likelihood that a word will be produced in response to a stimulus word. E.g. “chair” is the highest frequency response to the word “table” (Deese, 1959).

incoming stimuli (verbatim traces) and the meaning content of the stimuli (gist traces) in parallel, storing dissociated traces of verbatim and gist experience (Weekes et al., 2008). Verbatim traces represent episodically instantiated interpretations of surface forms, including contextual cues, whereas gist traces represent episodic interpretations of concepts (meanings, relations, and patterns) that are retrieved as a result of connecting the meaning across events (Weekes et al., 2008). FTT explains false recollection as a result of participants automatically inferring a theme from study words and then confusing studied with non-presented words that are good exemplars of the studied words.

In reading, it's known that children with comprehension difficulties have problems forming mental models of texts (e.g. Oakhill, 1996). The CI model (Kintsch, 1988; 1998) assumes that the representation of text occurs on three levels - the superficial, verbatim form of the text (the surface structure), and two subsequent levels that describe the mental representation of the text that includes bridging and elaborative inferences that go beyond what is stated explicitly (i.e. the text-base and the situation model).

Bransford and Franks (1971) found that when they presented participants with semantically related sentences, the verbatim forms of the sentences were rapidly forgotten. Sentences presented in a recognition test that fit with the 'gist meaning' of the sentences were falsely recognised as having been presented during the study phase with a high-degree of certainty. They explained this phenomenon by distinguishing between 'idea acquisition memory' and 'individual sentence memory', proposing that readers spontaneously integrate information expressed by related sentences, forming 'holistic, semantic ideas' (Branford & Franks, 1971).

This explanation fits with the fuzzy trace theory very well. Under this theory, verbatim forms of sentences decay rapidly, whilst the gist-based representation (i.e. the mental model) is relatively robust. As a consequence, the gist-based representation is then used as a template for recognition processes. It is proposed that these gist processes occur automatically as a result of activation spreading across associative connections between words (Reyna & Kiernan, 1994) or, in the Bransford and Franks experiment, between sentence propositions. There is evidence that readers with comprehension problems show deficiencies in detecting abstract semantic relations between words (e.g. Nation & Snowling, 1999). Presumably, these deficiencies extend to propositions at the sentence level as well as at the word-level.

Weekes et al. (2008) found evidence to suggest that children with comprehension problems have difficulties forming gist-based memory traces for semantically related words. In this study, children undertook phonological and semantic versions of the false memory paradigm described above. In the semantic version, poor comprehenders produced fewer non-studied related words (e.g. *sleep*) at recall and made fewer false-positive responses to non-presented but highly related words at recognition (e.g. *duvet*). In contrast, good comprehenders recalled greater numbers of highly related but non-presented words, and falsely recognised more of these words as having been presented during the learning phase. The two groups did not differ in their performances on the phonological version of the task, suggesting that these effects are specifically related to semantic processing.

With reference to Fuzzy Trace Theory (FTT) (Brainerd & Reyna, 2005), the fact that children with reading comprehension difficulties produced significantly fewer false memories suggests that the spreading of activation across semantic associations is inhibited or is in some way deficient in these individuals. With reference to the construction-integration model (Kintsch, 1988), children exhibiting specific comprehension deficits fail to activate related concepts or propositions occurring in the knowledge net, resulting in an impoverished text-base representation that precludes the integration of propositions into a coherent discourse representation or mental model of a text.

The present experiment sought to replicate these findings in an adult population. That is, participants divided into groups on the basis of reading comprehension performance should show different levels of false recall and recognition: Good comprehenders would show greater levels of false recollection in response to the semantic DRM than would poor comprehenders. The same effect, if based in semantic representations of words, would not occur in the phonological DRM.

3.2.2 Method

Participants

Sixty-nine adults participated in the experiment. The 54 women and 15 men received either £7.50 or course-credits for their participation. The participants' ages ranged from 18 to 45 years ($M = 21.07$, $SD = 6.04$). All participants were undergraduate or

postgraduate students at The University of Sussex. All spoke English as their native language and none were diagnosed as suffering from any reading or learning difficulties.

Materials and Procedure

The experiment was run over two sessions. The first session was run in groups or individually, depending on participant recruitment rates.

Session One

Reading Comprehension and Vocabulary

The Nelson-Denny Reading test (Brown, Fischco, & Hanna, 1993) was used to measure comprehension, vocabulary and reading rate.

The vocabulary component of the Nelson Denny Reading Test is comprised of eighty multiple-choice items, each with five response options. The words are drawn from high school and college text books and vary in difficulty. Participants were given a maximum of fifteen minutes to complete this component of the test.

The comprehension component of the Nelson Denny Reading Test requires participants to read seven passages and to answer a total of thirty eight questions, each with a choice of five answers. Participants were given twenty minutes to complete this part, the first minute being used to determine reading rate; participants were instructed to begin reading the first passage at their normal speed. After one minute had passed, the experimenter called 'mark'. At this point participants were instructed to draw a line underneath the sentence they were reading at that moment. The line number of this mark was then recorded by the experimenter once testing had finished.

If participants had not completed all of the questions within twenty minutes they were then given a 'bonus' ten minutes in which to do so (N.B. When correct responses were converted to standardised scores, scoring was reduced to accommodate this extra-time administration).

This test takes around 45 minutes to complete. Raw scores were transformed into standardised scores on the basis of the test manual.

Word-level tasks

Decoding skill was measured using 3 tasks obtained from Charles Perfetti and Leslie Hart. Two of these tasks require participants to detect phonologically legal yet orthographically incorrect words (i.e. pseudohomophones such as "EER"). In the first task, participants were given up to 10 minutes to circle up to 52 pseudohomophones, each one embedded in a choice of 3 non-word letter strings. They had 10 minutes to complete the task

The second task required participants to tick up to 37 pseudohomophones from a choice of 100 non-word letter strings in under 3 minutes.

The third task required participants to transform given words into orthographically legal new words through the deletion and addition of given phonemes, thereby giving a measure of phonological and orthographic skill together. Performance on this task was timed. Accuracy was divided by time to give a speed/accuracy measure.

For the purposes of statistical analysis, scores from the three tasks were transformed into z-scores, summed and divided by three. This composite measure of word-level ability is labelled 'word-level score' in the analyses below.

The False Memory Paradigm

Materials that are known to elicit false recollection in children aged from 9-11 years (i.e. those used in Holliday and Weekes, 2006 and Weekes et al., 2008) were used in the present experiment; the DRM paradigm (Deese, 1959; Roediger & McDermott, 1995). There were 2 versions of the task; a semantic version and a phonological version. In the semantic version, lists of words taken from Stadler, Roediger, & McDermott (1999) that were associated by meaning were used to create six study lists. These words were derived from association norms containing a critical word. The first words from the 6 Stadler et al. (1999) lists that appear in the top-half rank of false recollection of a critical word served as the study words.

Recognition tests were comprised of four studied words (targets) from each of the six lists, with targets taken from the 1st, 3rd, 5th, 7th and 9th positions. Also in the recognition test were three related but non-presented words for each list (one 'critical distracter' and 2 'related distracters') as well as two words unrelated in meaning or sound to the studied words for each list ('unrelated distracters').

In the phonological version of the paradigm, lists of words that are associated by sound were used in study and recollection trials. The lists were derived from norms containing a critical word and associated words that are related to that critical word by sound (initial phoneme, head or rime of the monosyllable) using computer-aided dictionary searches (see Westbury, Buchanan & Brown, 2002). Again, associated words were used to create six study lists. A recognition test comprised four studied words (targets) taken from the 1st, 3rd, 5th, 7th and 9th positions of each study list, as well as three

related words (1 ‘critical distracter’ and 2 ‘related distracters’) from each list, and two words unrelated in meaning or sound to the study words.

Participants were tested in groups or individually on the DRM paradigm, and took both the semantic and phonological versions. In both, six study lists of 10 items were presented at 4-second intervals via an audio CD (digitised using the Praat auditory manipulation package (Boresma, 2001)). All of these 10 items are associates of a critical, non-presented word. The order of presentation of the two versions of the paradigm was counterbalanced across participants.

Participants were instructed that they would be asked to remember a list of words. Presentation of each list was followed by a one-minute distracter task (digit substitution), after which they were asked to write down all those words that they were, “*sure* were presented in the study list.” This procedure was followed for all 6 lists.

A recognition test of 54 words presented in the experimenter’s voice was administered after all 6 lists had been presented. Participants were provided with an answer sheet and asked to circle ‘old’ for words presented in the study phase and ‘new’ for words not presented in the study phase.

Session Two

Session two was run individually on a participant-by-participant basis, usually within a week of participation in session one.

Working Memory

A sentence span task devised by Sarah McCallum (2002) was used to assess working memory. The task was comprised of 100 sentences arranged as five trials each in

sets of two, three, four, five and six sentences. These sentences were between eight and eleven words long. All were statements that could be either accepted (50) as making sense (e.g. “It is good exercise to run up and down the *stairs*.”), or rejected (50) as being nonsensical (e.g. “We can eat the honey that is singing to the *carpet*.”). Nonsensical statements were all created so as to be semantically or pragmatically incorrect – there were no syntactic violations. The sentences were presented orally and participants made verbal responses. The task required participants to firstly process the incoming information, and to store and then reproduce the last word from each sentence. Before testing began, participants were informed that the number of sentences in each set would increase as the test proceeded. Testing stopped once a participant failed to recall all stimuli in the correct order from three consecutive trials. Measures of accuracy of sentence verification (processing), and absolute memory span (the total number of correctly recalled words) were taken. Errors in sentence verification were extremely rare, but when they did occur, performance on the trial was discounted but was not recorded as an error trial. The sentence span task took from 5-10 minutes to administer.

Verbal Intelligence

Participants then took the National Adult Reading Test (NART) (Nelson, 1991). This test requires participants to read aloud a total of 50 irregular and infrequent words (e.g. ‘Quadruped’) and is used as a measure of verbal intelligence in clinical settings (e.g. Christiensen, Henderson, Griffiths, & Levings, 1997). None of the words used in the NART can be pronounced using regular grapheme-phoneme correspondences; all need to be read via the direct reading route. This route runs in the following manner (adapted from

Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001): First, the reader must have an entry for a given word in their mental lexicon. The features of the word's letters then activate the entry's letter units, these letter units then activate the word's entry in the orthographic lexicon and this word entry in the orthographic lexicon then activates the corresponding word entry in the phonological lexicon. The entry in the phonological lexicon must then activate the word's phonemes, triggering motor responses appropriate for a correct pronunciation (Coltheart et al., 2001). For these operations to run smoothly, the reader must have clearly specified entries at each stage, entries that are able to activate representations at subsequent stages coherently. A prerequisite to correct pronunciation therefore, is that the reader must be familiar with these words. Presumably, this familiarity stems from reading experiences and verbal communication. Thus, not only does the NART provide a measure of verbal intelligence, it also gives a rudimentary measure of a participants reading experience.

Items were presented on separate pieces of paper and participants were instructed to read-aloud each word. Before testing began, they were informed that the test was comprised of irregular words; many of which they would recognise and many that they may not. They were also informed that if they felt they had made an error they were allowed to read the word again. Responses were graded by the experimenter, and no feedback was given with regard to accuracy or performance. The total number of correctly pronounced words was used in the analyses that follow. This test took 5-10 minutes to administer.

3.2.3 Results

A total of 5 participants out of 69 stated that they had become aware of the purposes of the paradigm (i.e. the production of false memories) before the end of the session. These participants were excluded from the following analyses. Descriptive statistics for the entire sample are presented in Table 1.

Table 1. Descriptive Statistics (n = 64)

| | Mean | SD |
|----------------------------|--------|-------|
| ND Comprehension | 227.48 | 15.12 |
| ND Vocabulary | 236.11 | 14.94 |
| NART | 29.69 | 7.70 |
| Working Memory | 4.19 | 0.87 |
| Word-Level | -0.33 | 0.76 |
| Semantic Target Recall | 46.55 | 5.21 |
| Semantic CD Recall | 1.62 | 1.40 |
| Semantic RD Recall | 0.89 | 1.22 |
| Semantic UD Recall | 0.00 | 0.00 |
| Semantic Target A' | 0.96 | 0.07 |
| Semantic CD A' | 0.87 | 0.12 |
| Semantic RD A' | 0.62 | 0.14 |
| Phonological Target Recall | 35.97 | 9.00 |
| Phonological CD Recall | 0.09 | 0.29 |
| Phonological RD Recall | 2.02 | 2.00 |
| Phonological UD Recall | 0.28 | 0.68 |
| Phonological Target A' | 0.92 | 0.05 |
| Phonological CD A' | 0.73 | 0.11 |
| Phonological RD A' | 0.68 | 0.13 |

Grouping Variables and Characteristics

Participants were grouped into 'good' and 'poor' comprehenders on the basis of their performance on the comprehension subtest of the Nelson Denny Reading Test. The

poor group was comprised of those whose Nelson Denny comprehension scaled score fell below the group median of 232 ($N = 31$). The good comprehender group was comprised of those who scored at or above the whole group median of 232 ($N = 33$). The component reading skills of the two groups are summarised in Table 2. Significant differences (as assessed by independent samples t -tests) are highlighted with asterisks.

Table 2. Component reading characteristics of the two comprehension groups (mean(sd))

| | Reading Comprehension Group | |
|-------------------------------------|-----------------------------|------------------|
| | Good (n =33) | Poor (n = 31) |
| Mean ND comprehension scaled score: | 243.36 (9.59)** | 228.39 (15.83)** |
| Mean ND vocabulary scaled score: | 239.45 (5.11)** | 214.74 (11.26)** |
| Mean working memory: | 4.42 (0.97)* | 3.94 (0.68)* |
| Mean NART score: | 33.85 (6.86)** | 25.26 (5.91)** |
| Mean word-level score: | 0.33 (0.71)** | -0.42 (0.60)** |

** $p < .001$; * $p < .05$.

As shown in Table 2, the good comprehenders performed significantly better than the poor comprehenders on the comprehension component of the Nelson Denny Reading Test ($t(62) = 11.42, p < .001$) and the vocabulary component ($t(62) = 4.61, p < .001$). The good comprehenders had significantly larger working memory capacities than the poor comprehenders ($t(62) = 2.32, p < .01$). The good comprehenders also performed significantly better than the poor comprehenders on the National Adult Reading Test ($t(62) = 5.38, p < .001$), and on the word-level tasks ($t(62) = 4.59, p < .001$).

Group statistics for recall and recognition rates in both semantic and phonological versions of the DRM are presented in Figures 1, 2, 3 and 4.

Figure 1. Recall performance of good (n=33) and poor (n=31) comprehenders in the semantic DRM. Error bars represent one standard deviation above and below the mean

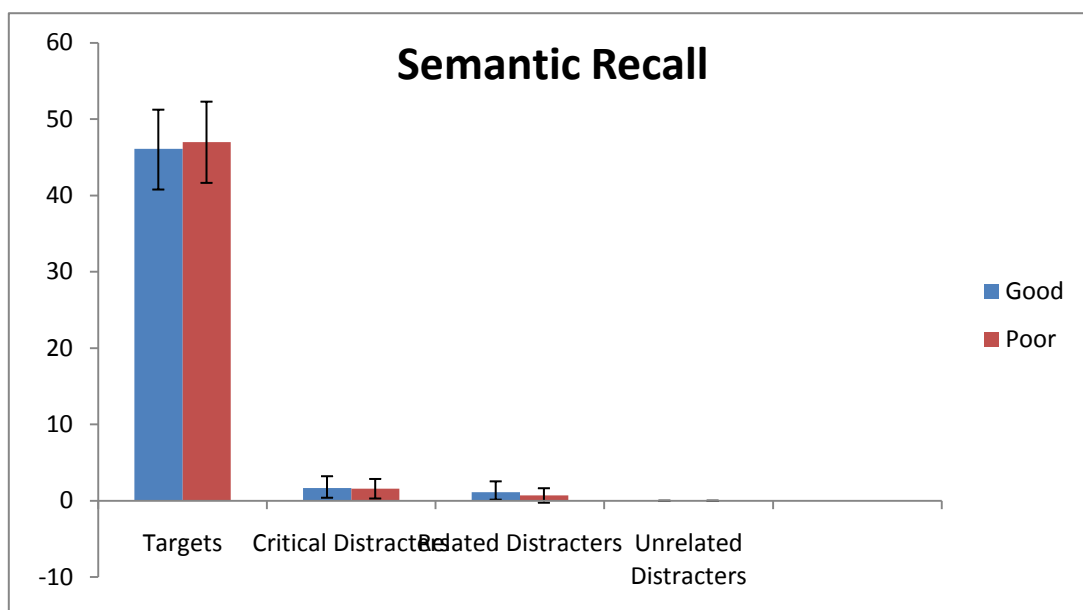
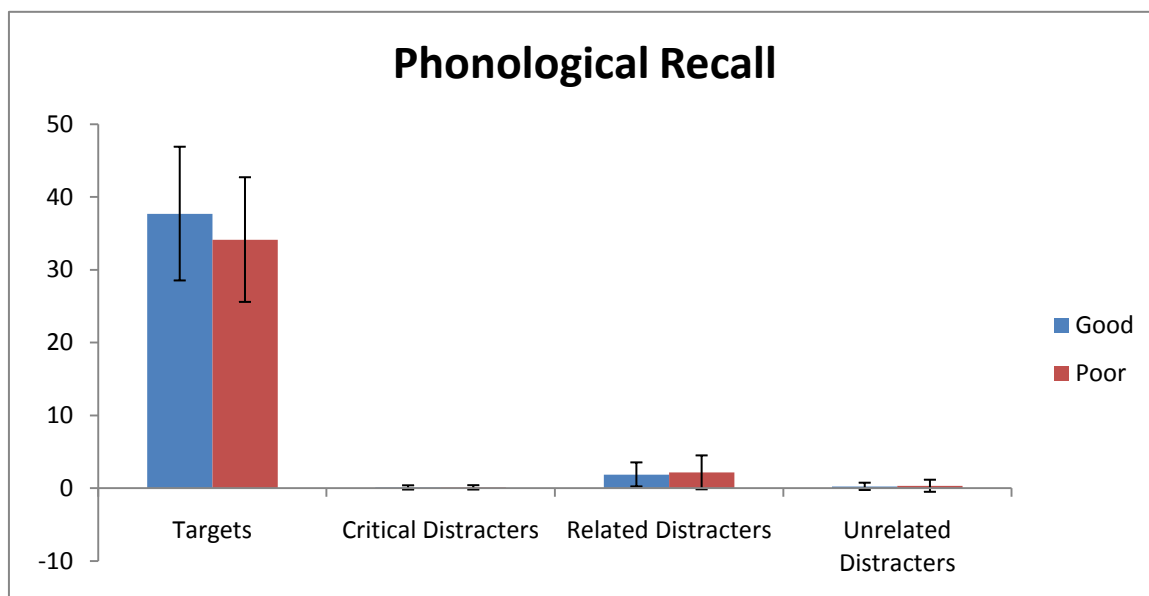


Figure 2. Recall performance of good (n=33) and poor (n=31) comprehenders in the phonological DRM. Error bars represent one standard deviation above and below the mean



A 2 x 2 mixed-design ANOVA was conducted on the recall data to determine whether there were differences in participants' levels of false memories. There were 2 within subjects variables (list type; item type), each with 2 levels (semantic DRM, phonological DRM; target, critical distracter). The between subjects factor was comprehension group.

There was a main effect of list type, $F(1, 62) = 179.72, p < .001$, indicating that more recollections (both veridical and false) were produced in response to semantically related word lists than to phonologically related word lists.

Contrary to predictions, there were no interactions with the between subjects variable comprehension group in any of the analyses: There was no interaction between comprehension group and list type ($F(1, 62) = 2.37, p = .13$); there was no interaction between comprehension group and item type ($F(1, 62) = 1.78, p = .19$); and there was no 3-way interaction between list type, item type and comprehension group ($F(1, 62) = 2.92, p = .14$).

Another 2 x 2 ANOVA was conducted, this time including 'related distracter' recall as an independent variable. Again, there were 2 within subject variables (list type; item type), each with 2 levels (semantic DRM, phonological DRM; target, related distracter). The between subjects variable was comprehension group.

There was a main effect of list type, ($F(1, 62) = 125.13, p < .001$). As before, there were no interactions with the between subjects factor comprehension group in any of the analyses: There was no interaction between comprehension group and list type ($F(1, 62) = 2.74, p = .10$); there was no interaction between comprehension group and item type ($F(1,$

62) = 2.20, $p = .14$); and there was no 3-way interaction between list type, item type and comprehension group ($F(1, 62) = 1.87$, $p = .18$).

Recognition

As recommended by Snodgrass and Corwin (1988), recognition data were transformed into A' values to discriminate signal from noise (i.e. answers to targets and distracters are corrected for 'yes saying' on the basis of false-positive responses to unrelated distracters). The formula proposed by Stanislaw and Todorov (1999) was used for this transformation (see Appendix 1).

Figure 3. Recognition A' values for good ($n=33$) and poor ($n=31$) comprehenders in the semantic version of the DRM paradigm

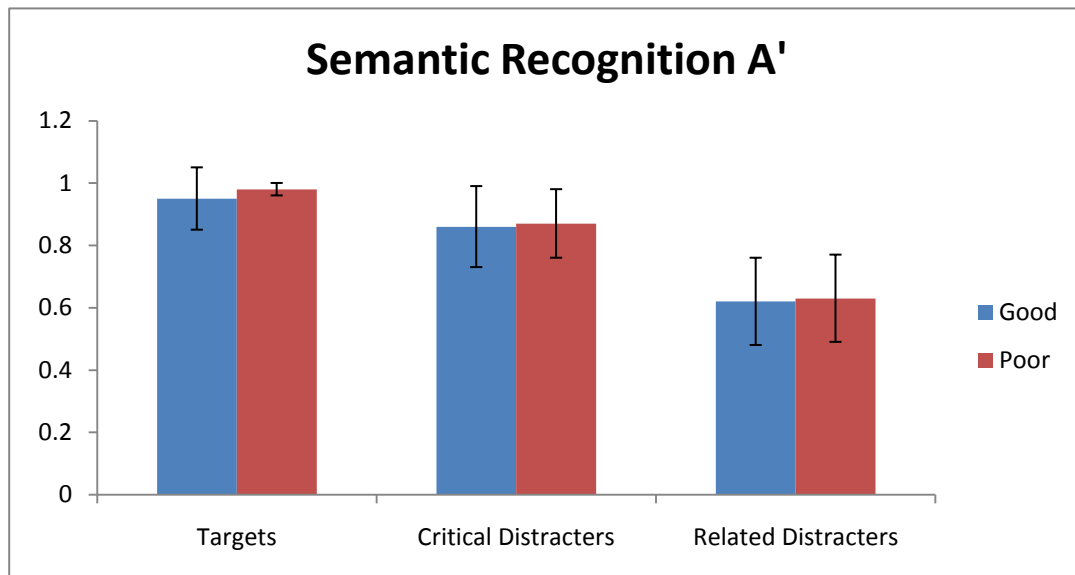
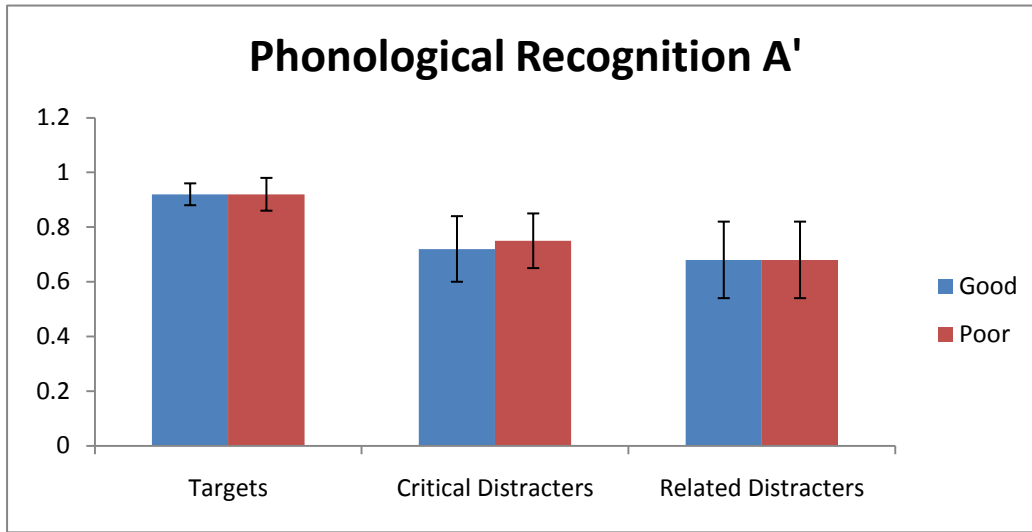


Figure 4. Recognition A' values for good (n=33) and poor (n=31) comprehenders in the phonological version of the DRM paradigm



A 2 x 2 mixed-design ANOVA was conducted on the recognition data to determine whether there were differences in participants' levels of false recognition. There were 2 within subjects variables (list type; item type), each with 2 levels (semantic DRM, phonological DRM; target, related distracter). The between subjects variable was comprehension group.

Again, there was a main effect of list type ($F(1, 62) = 36.21, p < .001$), indicating that more recognitions (both veridical and false) were produced in response to semantically related word lists than to phonologically related word lists. There was a main effect of item type ($F(1, 62) = 195.73, p < .001$), suggesting that participants recognised (proportionally) more targets than they did non-presented critical distracters.

Contrary to predictions, there were no main effects or interactions with the between-subjects variable comprehension group: There was no interaction between list

type and comprehension group ($F(1, 62) = 0.02, p = .88$); there was no interaction between item type and comprehension group ($F(1, 62) = 0.02, p = .89$); and there was no 3-way interaction between list type, item type and comprehension group ($F(1, 62) = 1.44, p = .23$).

Another 2 x 2 ANOVA was conducted on the recognition data, this time with 'related distracter' recognition as an independent variable. Again, there were 2 within subject variables (list type; item type), each with 2 levels (semantic DRM, phonological DRM; target, related distracter). The between subjects factor was comprehension group.

There was no main effect of list type ($F(1, 62) = 0.32, p = .58$), suggesting that recognition rates (regardless of whether they were veridical or false) did not differ between the two types of lists. There was a main effect of item type ($F(1, 62) = 450.73, p < .001$), suggesting that participants correctly recognised more targets than related distracters. There was no interaction between list type and comprehension group ($F(1, 62) = 0.16, p = .69$). There was no interaction between item type and comprehension group ($F(1, 62) = 0.17, p = .69$). There was no 3-way interaction between list type, item type and comprehension group ($F(1, 62) = 0.07, p = .80$).

3.2.4 Conclusions

Contrary to predictions, there were no differences in false recollection between good and poor comprehenders. One potential explanation of the non-replication of Weekes et al. (2008) is that the children in the Weekes et al. (2008) study were at an age (9-11 years) at which their lexical representations were relatively underspecified; it may

be that the poor comprehenders had lexical representations that were relatively poorly specified (and thereby poorly 'networked' to their associates) as to be relatively impervious to the DRM paradigm, at least in comparison to the good comprehenders. By adulthood, these lexical representations are too robust in both good and poor comprehenders to show any real differences between them. So, it could be that there is some critical period in which the relation between gist-memory and reading comprehension can be assessed using the paradigm.

An alternative hypothesis is that the participants in the present study did not reflect the true range of reading comprehension ability expressed in the general population. The poor comprehenders in the present study were only classified as such on the basis of mean scores on the comprehension component of the Nelson Denny Reading Test falling below the group median of 232. In the study by Weekes and colleagues (2008), the poor comprehenders were those whose comprehension age (as determined by norms for the Neale Analysis of Reading Ability) fell at least one year below their reading accuracy age and one year below their chronological age (as in Oakhill, 1982). The poor comprehender group in this study had reading problems specific to comprehension rather than more general reading problems. It could be, therefore, that the readers classified as 'poor' in the present study were in fact 'weaker' rather than poor comprehenders. As shown in Table 1, the two groups were not matched, as they were in the Weekes et al. (2008) study, for vocabulary or non-word reading (i.e. word-level task performances). It could be said that these readers were weaker because of lower-level processing difficulties (i.e. decoding, vocabulary, working memory) rather than specific problems with

gist-based memory processing. However, if this were true then we would expect the weaker group to have performed more poorly than the stronger group in their recall and recognition rates of the target items, which was not the case.

Long and Chong (2001) have found evidence to suggest that adults with reading comprehension problems activate (and reactivate) information from earlier discourse to the same extent as do good comprehenders. In this study, good and poor comprehenders were presented with stories in which a character's action was either consistent or inconsistent with a description of the character presented earlier in the story: The story's protagonist (Mary) is described as either a strict vegetarian or as a fast food addict. Later in the text, the statement, "Mary ordered a cheeseburger and fries," was presented; a statement that is either consistent or inconsistent with the character description. When the character description was presented in text adjacent to the inconsistent statement (local coherence), both groups of readers showed an inconsistency effect (i.e. longer reading times for the statement). When the character description and the inconsistent statement were separated by intervening text (global coherence), only the good comprehenders showed an inconsistency effect. This finding is in line with the proposal that poor comprehenders have trouble maintain global coherence during reading. Their 2nd experiment aimed to investigate whether the findings of Experiment 1 were attributable to poor comprehenders failing to re-activate prior text information. In this experiment, subjects were presented with probes about the character descriptions at three positions: (a) after the character description; (b) after the intervening (filler) sentences; and (c) after the consistent/inconsistent statement. Probe verification latencies

did not differ between the two groups at any time point, suggesting that both groups of readers re-activated prior text concepts when the text required it. Thus, good and poor adult comprehenders do not differ in their ability to activate information that is gleaned from texts, consistent with the present finding that good and poor comprehenders show similar patterns of false recall and recognition in the DRM paradigm. They do differ in their ability to integrate this information across sentences, however. Specifically, when the to-be-reactivated information is separated by intervening text, the conflict statement fails to produce a conflict effect in poor comprehenders. With reference to the CI model, poor comprehenders do not show difficulties in *construction*, but they do show difficulties in *integration*.

The next chapter of this thesis sought to measure the component skills that may be related to reading comprehension skill: Decoding ability, verbal intelligence, vocabulary and working memory capacity. The fifth chapter sought to measure abilities that may be pertinent to the integration of information across sentence boundaries in a conflict paradigm that manipulated coherence and cohesion, with the proposal that working memory and/or semantic ability may be related to the integration process.

4 Component Skills of Comprehension

4.1 Chapter Overview

The experiment of Chapter 3 suggested that adults with comprehension weaknesses do not have difficulties activating semantically related concepts in the DRM paradigm. The conclusion of this chapter was that these readers may have difficulties integrating related concepts to form an accurate text-base and subsequent discourse-level of textual representation.

In Chapter 2 (Literature Review) a number of models of reading comprehension were presented and described. The CI model (Kintsch, 1988; 1998) assumes that linguistic input serves to activate related concepts in readers' knowledge, implying that words that are encountered during reading need to be adequately represented in the mental lexicon. The Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) is specific in its contention that inferior lexical codes impact on reading directly (contending that syntactic analysis will be constrained if semantic specificity cannot be established) and indirectly (by inefficient lexical access consuming resources that would otherwise be devoted to higher-order comprehension processes).

The capacity-constrained model of comprehension (Just & Carpenter, 1992) proposes that comprehension difficulties can be attributed to constraints in working memory capacity. On the other hand, the long-term working memory (Ericsson & Kintsch, 1995) and the connectionist account of working memory (MacDonald & Christiansen,

2002) both attribute differences in comprehension skill to differences in linguistic experience and proficiency rather than working memory capacity per se.

In the present study, measures of word-level ability, vocabulary, working memory capacity and verbal intelligence (a rudimentary measure of linguistic experience) were taken, and their ability to predict variance in performance on the comprehension component of the Nelson Denny Reading Test was assessed through correlational and multiple regression analyses. The results of these analyses implicate vocabulary, word-level ability and, to some degree, verbal intelligence as being the strongest predictors of reading comprehension.

The discussion focuses on the partial support for the Lexical Quality Hypothesis provided by the predictive power of word level performance to variance in comprehension skill. Also discussed is the modest support for the connectionist and long-term working memory accounts of reading comprehension that is also implied by the predictive efficacy of word-level performance, vocabulary and verbal intelligence. The shortcomings of the Capacity-constrained model of working memory and comprehension skill (Just & carpenter, 1992) is discussed in light of the fact that performances on the sentence span task failed to predict variance in comprehension skill once the contributions of other variables had been accounted for.

4.2 Experiment 2: Word-Level Ability, Vocabulary, and Working Memory as Predictors of Reading Comprehension Skill

4.2.1 Introduction

When reading a text, comprehension occurs as the reader builds a mental representation of the text's message. This constructed mental representation can be conceptualised as a mental image of a scene extracted from a text and, in the current literature, is referred to as the *discourse or situation model*' of a given text (e.g. Kintsch, 1988). A coherent discourse model should include inferences that impart information that goes beyond the literal, concrete representations obtained from actual sentences. Inherently, such inferences rely upon world or semantic knowledge for their elucidation.

Poor comprehension is characterised by a failure to make appropriate inferences during reading (e.g. Baker & Stein, 1981), including problems integrating general knowledge with information in a given text to fill in missing details (Cain & Oakhill, 1999), as well as integrating ideas in a text in order to answer questions about main ideas and themes (Yuill & Oakhill, 1991). At word and sentence levels, poor comprehenders have slower and less efficient word-identification skills and greater difficulty processing low-frequency syntactic structures (Bell & Perfetti, 1994). These problems appear to be present both in adults and children demonstrating reading comprehension impairments (Long et al., 2006)

Kintsch's CI model (1988; 1998) describes how linguistic input activates semantic knowledge which is subsequently integrated to form a coherent mental representation of

a text. First, the surface structure encodes the verbatim form of the text, thereby providing syntactic information and orthographic/phonological representations for specific lexical items. A semantic representation of the text is derived as a function of this surface structure by linking the reader's background knowledge to the propositions given by the text, leading to the construction of a text base representation. This background knowledge is conceptualised as an associative network comprised of inhibitory and excitatory links between nodes of concepts or propositions (Kintsch, 1988).

Linguistic input therefore serves to activate networks comprising nodes of concepts or propositions within the knowledge net, with sufficiently activated nodes forming a separate network from which a text-base is derived (the integration phase). This text base is then used to construct a coherent discourse-level representation of the text via the integration of sentence level propositions and inferences.

Word-Level Influences

As stated above, the construction of the text base depends upon accurate representations of lexical items. The so-called Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) holds that the higher-order process of reading comprehension relies upon a foundational processing factor: the effectiveness of basic word identification skill (Perfetti & Hart, 2002). Thus, reading comprehension problems, under this hypothesis, result from impoverished lexical representations that constrain higher-order comprehension processes.

A high-quality lexical code should have orthographic phonological and semantic specificity, allowing a given word to be understood (and therefore encoded) *coherently*

when it is encountered in both auditory and visual modalities (Perfetti, 1985). Coherence, in this case, refers to the idea that lexical codes (for phonology, orthography and semantics) are accessed in synchrony, giving the impression of a unitary word perception event (Perfetti & Hart, 2002).

As stated in Chapter 2, inferior lexical codes affect reading comprehension by: hindering the production of an accurate text base; inhibiting the activation of semantic associates and concepts; and consuming resources that would otherwise be devoted to higher-order comprehension processes.

Working Memory

Working memory has been conceptualised as a limited capacity temporary space used for the storage and manipulation of incoming information (e.g. Baddeley & Hitch, 1974). Text comprehension requires the use of executive functions such as structuring, monitoring and problem solving (Ferstl, Guthke & von Cramon, 2002). As such, working memory is said to be responsible for the integration of ideas across sentences and acts as an interface between individual sentence meanings and knowledge domains in order to make inferences (leading to the construction of an accurate discourse representation). As working memory capacity is limited (e.g. Daneman & Carpenter, 1980), individual differences in reading comprehension skill are believed to result from individual differences in working memory capacity (e.g. Just & Carpenter's (1992) Capacity Theory of comprehension).

Alternatively, individual differences in working memory (and comprehension) may be attributed to individual differences in skill and knowledge rather than differences in

capacity. That is, it is the *functioning* of WM that differs between good and poor comprehenders. The connectionist-based account of variation in working memory (*The distributed learning model* - MacDonald & Christiansen, 2002), proposes that capacity is a property of the functional architecture of working memory (i.e. the number of processing units, how activation passes through weights, etc.), and that this functioning is modulated by the network's experience (i.e. how often it has processed similar input in the past) (Long et al., 2006).

Another theory of working memory proposes that a single, unitary and capacity-constrained conception of working memory is too restrictive to explain the complex and cognitively demanding task of reading comprehension. Some have proposed that working memory comprises a limited capacity focus of attention that works in unison with aspects of long term memory that are accessed via the activation of retrieval cues. Ericsson & Kintsch (1995) propose a model of working memory that consists of the standard limited-capacity mechanism that they call *short-term working memory*, and a mechanism based on skilled storage and retrieval in long-term memory that they call *long-term working memory*. Like the connectionist account, this model emphasises the importance of practice and skill in reading comprehension, rather than capacity limitations. According to this model, reading comprehension skill should vary as a function of reading experience and the depth or breadth of the reader's domain or semantic knowledge of the topics given in a text.

In the present experiment, a number of measures thought to contribute to reading comprehension ability were taken, and were entered into a regression model as

predictors of performance on the comprehension component of the Nelson Denny Reading Test. Word-level ability and vocabulary were measured with two decoding tasks, an orthographic transformation task and the vocabulary component of the Nelson Denny Reading Test. Working memory capacity was measured with a variant of Daneman and Carpenter's (1980) sentence span task. Verbal intelligence was taken as a measure of participants' overall language ability derived from linguistic experience.

The Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) would predict that variance in both word-level ability and working capacity would each predict unique variance in comprehension skill. In contrast, the distributed learning account of working memory (MacDonald & Christiansen, 2002) would predict that only the measures of linguistic competence (word-level ability, vocabulary and verbal intelligence) would predict variance in comprehension skill, with performance on the sentence span task predicting none of the variance when these variables were accounted for. The long-term working memory model (Ericsson & Kintsch, 1995) attributes variance in comprehension skill to differences in access to world knowledge. As such, measures of word knowledge (vocabulary and verbal intelligence as measured by the NART) should each predict unique variance in Nelson Denny comprehension scores, over and above that predicted by performances on the sentence span task.

4.2.2 Method

Participants

The data for Experiments 1 and 2 were collected simultaneously. Thus, the same sixty-nine adults who participated in Experiment 1 also participated in the present experiment. The 54 women and 15 men received either £7.50 or course-credits for their participation. The participants' ages ranged from 18 to 45 years ($M = 21.07$, $SD = 6.04$). All participants were undergraduate or postgraduate students at The University of Sussex. All spoke English as their native language and none were diagnosed as suffering from any reading or learning difficulties.

Materials and Procedure

Word-Level Tasks

As in Experiment 1, decoding skill was measured using 3 tasks obtained from Charles Perfetti and Leslie Hart. Two of these tasks require participants to detect phonologically legal yet orthographically incorrect words (i.e. pseudohomophones such as "EER"). In the first task, participants were given up to 10 minutes to circle up to 52 pseudohomophones, each one embedded in a choice of 3 non-word letter strings. They had 10 minutes to complete the task. This measure is labelled, 'Pseudohomophone/3', in the analyses presented below.

The second task required participants to tick up to 37 pseudohomophones from a choice of 100 non-word letter strings in under 3 minutes. This is labelled as 'Pseudohomophone 3 minutes' in the analyses below.

The third task required participants to transform given words into orthographically legal new words through the deletion and addition of given phonemes, thereby giving a measure of phonological and orthographic skill together. Performance on this task was timed. Accuracy was divided by time to give a speed/accuracy measure. This measure is labelled 'Transformation' in the analyses below.

Vocabulary and Comprehension

The Nelson-Denny Reading test (Brown, Fischco, & Hanna, 1993) was used to measure comprehension, vocabulary and reading rate. This test takes around 45 minutes to complete. Test characteristics and procedure were as described in Experiment 1.

Verbal Intelligence

The National Adult Reading Test (Nelson, 1991) was used to assess verbal intelligence. Test characteristics and administration procedures are as described in Experiment 1.

Working Memory

The sentence span task devised by Sarah McCallum (2002) was used to assess working memory capacity. Test characteristics and administration procedures are as described in Experiment 1.

4.2.3 Results

Table 3. Descriptive Statistics (n=69)

| | Mean | SD |
|---------------------------|--------|-------|
| ND Comprehension | 228.14 | 15.20 |
| ND Vocabulary | 236.49 | 14.68 |
| NART | 29.84 | 7.62 |
| Working Memory | 4.19 | 0.90 |
| Transformation task | 0.09 | 0.05 |
| Pseudohomophone/3 | 43.57 | 8.51 |
| Pseudohomophone 3 minutes | 24.14 | 6.86 |

For an initial assessment of the relationships between each of the measures, correlation coefficients between them were calculated. These analyses are presented in Table 4.

Table 4. Spearman's rho correlation coefficients (n=69)

| | Comprehension | Vocabulary | NART | WM | Trans. task | PsH/3 | PsH/3 minutes |
|---------------------------|---------------|------------|--------|--------|----------------|--------|------------------|
| ND Comprehension | 1.00 | | | | | | |
| ND Vocabulary | .567** | 1.00 | | | | | |
| NART | .598** | .787** | 1.00 | | | | |
| Working Memory | .294* | .316** | .341** | 1.00 | | | |
| Transformation task | .372** | .466** | .533** | .238* | 1.00 | | |
| Pseudohomophone/3 | .373** | .401** | .461** | .360** | .362** | 1.00 | |
| Pseudohomophone 3 minutes | .529** | .594** | .629** | .401** | .561** | .443** | 1.00 |

**p<.001; *p<.05.

As shown in Table 4, all of the measures correlate significantly with each other, and with performances on the comprehension component of the Nelson Denny Reading Test. Since the coefficients between the decoding and transformation tasks are all relatively high, these values were transformed into z-scores, summed and then averaged.

Z scores over the three tasks were used to provide a composite measure of word-level skills. This variable is hereafter referred to as 'word-level'. The correlation coefficients are presented in Table 5.

Table 5. Spearman's rho correlation coefficients (n=69)

| | ND Comprehension | Vocabulary | NART | Working memory | Word-level |
|------------------|---------------------|------------|--------|-------------------|------------|
| ND Comprehension | 1.00 | | | | |
| Vocabulary | .567** | 1.00 | | | |
| NART | .598** | .787** | 1.00 | | |
| Working Memory | .294* | .316** | .341** | 1.00 | |
| Word-level | .499** | .558** | .626** | .417** | 1.00 |

** $p < .001$; * $p < .05$.

To examine the relative contributions of each of the measures variables in predicting Nelson Denny comprehension performance, the data were entered into a stepwise, hierarchical multiple regression analysis. The order in which each variable was entered into the model was determined by the strength of its association with ND comprehension (i.e. their correlation coefficients), from strongest to weakest. This analysis is presented in Table 6.

Table 6. Stepwise hierarchical multiple regression analysis, with ND comprehension as the dependent variable (n=69)

| | <i>B</i> | <i>SE B</i> | β | ΔR^2 |
|----------------|----------|-------------|---------|--------------|
| Step 1 | | | | |
| Constant | 195.97 | 6.31 | | |
| NART | 1.08 | .21 | .54*** | .29*** |
| Step 2 | | | | |
| Constant | 156.94 | 29.81 | | |
| NART | .81 | .29 | .41** | |
| Vocabulary | .20 | .15 | .19 | .02 |
| Step 3 | | | | |
| Constant | 173.20 | 30.70 | | |
| NART | .55 | .32 | .28^ | |
| Vocabulary | .16 | .15 | .16 | |
| Word-level | 4.71 | 2.63 | .24^ | .03^ |
| Step 4 | | | | |
| Constant | 171.61 | 32.06 | | |
| NART | .54 | .32 | .27^ | |
| Vocabulary | .16 | .15 | .16 | |
| Word-level | 4.54 | 2.79 | .23 | |
| Working Memory | .36 | 1.94 | .02 | .00 |

*** $p < .001$; ** $p < .01$; * $p < .05$; ^ $p < .10$.

Given the significant correlations presented in Table 5, it comes as some surprise that the only significant predictor of Nelson Denny comprehension score is performance on the National Adult Reading Test. As stated in the materials section, this test gives a rudimentary assessment of reading experience. It may, therefore, represent a composite variable comprised of many factors, some of which may be directly related to reading skill and some of which may represent more general cognitive factors whose relation to

reading comprehension is relatively unknown. Thus, from a theory-driven perspective, verbal IQ warrants a place in the predictive model *after* rather than before other more specific and well-defined variables (Catts, Fey, Zhang and Tomblin, 1999). As can be seen in Table 6, entering this variable into the model first partials out the variance accounted for by the other variables, whose relation to reading has been empirically established. In keeping with the method employed by Catts et al. (1999), it was decided that raw scores on the National Adult Reading Test should be entered into the regression model after the theory-based measures. This analysis is presented in Table 7.

Table 7. Stepwise hierarchical multiple regression analysis, with ND comprehension as the dependent variable (n=69)

| | <i>B</i> | <i>SE B</i> | β | ΔR^2 |
|----------------|----------|-------------|---------|--------------|
| Step 1 | | | | |
| Constant | 111.33 | 26.34 | | |
| Vocabulary | .49 | .11 | .48*** | .23*** |
| Step 2 | | | | |
| Constant | 156.06 | 29.49 | | |
| Vocabulary | .31 | .13 | .29* | |
| Word-level | 6.80 | 2.37 | .35** | .09** |
| Step 3 | | | | |
| Constant | 153.44 | 30.63 | | |
| Vocabulary | .30 | .13 | .29* | |
| Word-level | 6.45 | 2.59 | .33* | |
| Working Memory | .68 | 1.96 | .04 | .00 |
| Step 4 | | | | |
| Constant | 171.61 | 32.06 | | |
| Vocabulary | .16 | .15 | .16 | |
| Word-level | 4.54 | 2.79 | .23 | |
| Working Memory | .36 | 1.94 | .02 | |
| NART | .54 | .32 | .27^ | .03^ |

*** $p < .001$; ** $p < .01$; * $p < .05$; ^ $p < .10$.

As shown in Table 7, the only significant predictors of variance in Nelson Denny comprehension scores are vocabulary and the composite word-level variable (accounting for 23% and 9% of the variance respectively). Surprisingly, working memory does not account for any significant amount of variance in Nelson Denny comprehension scores. When entered into the model last, the National Adult Reading Test accounts for 3% of the variance in Nelson Denny comprehension scores, which is marginally significant.

4.2.4 Conclusions

To some extent, the analysis presented in Table 7 accords with the lexical quality hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) presented in the introduction. This theory holds that the higher-order process of reading comprehension relies upon the effectiveness of basic word identification skill. When this process is deficient or laboured, working memory processes are consumed, monopolising resources that would otherwise be devoted to higher-order comprehension processes. The finding that vocabulary and word-level skills both make substantial and unique contributions to reading comprehension would seem to accord with this proposal. However, if the lexical quality hypothesis is correct in its assertion that working memory capacity or processing is *fundamental* to reading comprehension skill, then the fact that working memory does not account for **any** of the variance in this measure is disquieting.

The perspectives that best explain the present data is the long-term-working-memory model of Ericsson and Kintsch (1995) and/or the distributed learning model of MacDonald and Christiansen (2002). According to these perspectives, it is variance in reading *experience* that determines readers' comprehension abilities. In simple terms, an experienced reader has been exposed to text to a greater extent than has an inexperienced reader. Prolonged exposure to text increases the proficiency with which one can access words in the lexicon (i.e. vocabulary skill) and the ability to decode new or unfamiliar words (i.e. word-level skills). The fact that these two variables are the only statistically significant predictors of variance in comprehension skill (among the measures

that were taken) supports the hypothesis that it is the functioning of working memory rather than capacity that determines reading comprehension skill.

The question remains as to whether these skills are determined by reading experience itself or whether good linguistic ability drives people to engage in more reading. As stated in the materials section, The National Adult Reading Test gives a rudimentary measure of verbal intelligence and language experience. That fact that this measure accounts for 3% of the variance in comprehension skill (a value that approaches statistical significance), when the variance predicted by the other measures has been accounted for, implies that reading experience does indeed have some effect on comprehension skill. Whether this effect is driven by the readers' knowledge with regard to the topics given by the test materials (through multiple experiences with narrative texts), or by more sophisticated syntactic awareness or ability (derived from text experience more generally), cannot be resolved with the present data.

By requiring participants to answer comprehension questions after they have read each passage, the comprehension component of the Nelson Denny Reading Test gives a measure of off-line comprehension skill. Experiments 3 and 4 sought to reveal the linguistic and cognitive skills that are related to on-line reading comprehension. In both experiments, participants read two-sentence stories in which coherence (whether the two sentences made sense together) and cohesion (the presence or absence of cohesive ties) were manipulated. In Experiment 3, measures of decoding skill, vocabulary and working memory capacity were taken, and their relations to performance on the coherence and cohesion paradigm were assessed with regression and ANOVA analyses. In Experiment 4,

verbal intelligence and several measures of semantic access were taken in addition to the measures of Experiment 3.

5 Component Skills, Semantic Access and Judgements of Coherence

5.1 Chapter Overview

Chapter 4 showed that word-level ability and vocabulary both account for unique portions of variance in comprehension skill, a finding that provides partial support for the lexical quality hypothesis (Perfetti & Hart, 2002; Perfetti, 2007). Full support for the hypothesis would have been obtained if working memory capacity uniquely accounted for a significant portion of the variance in comprehension skill. This was not the case when comprehension skill was measured off-line.

The experiments presented in this chapter seek to illuminate the processes and skills that are important to on-line text processing and comprehension. A coherence judgement task developed by Ferstl and colleagues (Ferstl & von Cramon, 2001; Ferstl, Guthke, & von Cramon, 2002; Ferstl, 2006) is used for this purpose. Four conditions that result from crossing coherence and cohesion (i.e. the presence of a lexical connection) are used: Coherent and cohesive (where sentences cohere, and a cohesive tie makes their coherence explicit); coherent and incohesive (where sentences cohere, but coherence has to be inferred on the basis of pragmatic information rather than lexical cohesion); incoherent and cohesive (where sentences that do not cohere are erroneously linked with a cohesive tie); and incoherent and incohesive (where sentences do not cohere, and are not erroneously linked with a cohesive tie). Typically, the paradigm elicits an interaction between coherence and cohesion in reading times for the second (target) sentence:

Targets in coherent and cohesive trials are read more quickly than targets in coherent and incohesive trials; and targets in incoherent and incohesive trials are read more quickly than are targets in incoherent and cohesive trials. Both of the experiments in this chapter demonstrate these interactions in both the error and the standardised reading time data. In Experiment 3, variance in the size of the reading time interaction is predicted by working memory capacity. In Experiment 4, this variance is predicted by semantic fluency rather than working memory capacity. None of the variance in the magnitude of this interaction effect was predicted by any of the word-level measures (vocabulary, decoding ability or verbal intelligence), and therefore did not provide any evidence in favour of the Lexical Quality Hypothesis (Perfetti & Hart, 2002), or the Distributed Learning Model (MacDonald and Christiansen (2002). The results are interpreted with reference to the capacity constrained model of working memory and reading comprehension skill (Just and Carpenter, 1992), the long term working memory model (Ericsson & Kintsch, 1995), comprehension monitoring, and semantic ability.

5.2 General Introduction

When readers understand a text and construct a situation model of the text, their representations need to be coherent at both *local* and *global* levels. Local coherence involves mapping each proposition in an incoming sentence to other propositions currently active in working memory, usually propositions from the immediately preceding sentence or two. Global coherence involves mapping incoming propositions to

information encountered earlier in the text (beyond the span of working memory) and to relevant world knowledge (Long & Chong, 2001).

Global coherence is involved in the development of a coherent situation model of a text. A situation model is a mental representation of the people, setting, actions, and events that are mentioned in explicit clauses or are inferred on the basis of world knowledge (Graesser, Singer, & Trabasso, 1994). These knowledge-based inferences can depict a vast array of situations and events; they may include the goals and plans that motivate a particular character (including representations of that character's knowledge and beliefs, their traits and emotions), properties of particular objects, spatial relationships among entities, referents of nouns and pronouns and so on. Some of these inferences are generated during the course of comprehension (i.e. on-line), whereas others are normally generated when the reader recollects the story at a later date (i.e. off-line) (Graesser et al., 1994).

The only measure of reading comprehension that has been used in this thesis so far has been the comprehension component of the Nelson Denny Reading Test. As described in Experiment 1, this test requires participants to read a passage and to then answer questions about it. Presumably, these comprehension questions act as retrieval cues for the text, so it could be said that this test provides a measure of off-line comprehension processing.

The present studies seek to reveal which component reading skills are related to successful on-line comprehension; skills related to the generation of on-line inferences. McKoon and Ratcliff (1992) propose that only bridging inferences required to establish

local coherence are generated on-line. According to their 'minimalist hypothesis', only those inferences that can be made on the basis of easily accessible information (either from explicit statements in the text or from general knowledge) are generated automatically during the reading process. These representations of textual information then provide the basis for more goal-directed, purposeful global inferences at a later date (McKoon & Ratcliff, 1992). Others argue (e.g. Trabasso & van den Broek, 1985) that causal and goal-related inferences are made on-line during reading. These authors propose that the representation of *why* certain events occur and *why* actions are carried-out is one of the main functions of comprehension. Trabasso and van den Broek (1985) analysed immediate and delayed recall data for narrative texts from a number of studies, and found that memory for story events was most pronounced for events that were part of a causal chain and had many causal connections to other events in the stories. Similarly, O'Brien and Albrecht (Albrecht & O'Brien, 1993; O'Brien & Albrecht, 1992) conducted reading time studies to show that global inconsistencies in narrative texts can affect processing time, even when local coherence is maintained. This suggests that comprehension processes that serve to create a global representation of a text may occur on-line during the reading process itself, rather than as part of a memory retrieval process.

The studies described above used texts that varied in length and in their degree of local and global coherence. Rather than measure participants' memories and reading times for natural texts, the present studies sought to measure reading times for 'minimal stories' in which descriptions of why events occurred and why actions were carried-out had to be inferred on the basis of world knowledge. A sentence comprehension task

developed by Ferstl and colleagues (Ferstl & von Cramon, 2001; Ferstl, Guthke, & von Cramon, 2002; Ferstl, 2006) was used for this purpose. The materials comprised 96 sentence pairs about everyday situations from a wide variety of topics. Each pair comprised a context sentence and a target sentence that either made sense together (the coherent condition) or did not make sense together (the incoherent condition). Each target sentence occurred in four versions: the cohesive version contained one or two lexical items (e.g. pronouns or conjunctions) that explicitly signalled the connection between the sentences. In the incohesive version, these cohesive ties were omitted or replaced so that the relationship between the two sentences had to be inferred on the basis of pragmatic rather than lexical information. The incoherent conditions were created by switching the context sentences of two coherent trials. As was the case for the coherent conditions, the target sentences in the incoherent conditions appeared both in cohesive and incohesive conditions. The paradigm used a 2 x 2 within-subjects design, with the variables coherence and cohesion: 'coherent and cohesive' (yes or no); 'coherent and incohesive' (yes or no); 'incoherent and cohesive' (yes or no); and 'incoherent and incohesive' (yes or no) (n.b. Coherent trials should always be answered 'yes' (regardless of cohesion), and incoherent trials should always be answered 'no'). Examples of the four conditions are shown in Table 8.

Table 8. Example materials from the four conditions of the Coherence and Cohesion paradigm (adapted from Ferstl et al., 2002)

| | Coherent | Incoherent |
|-------------------|--|---|
| Cohesive | Mary's exam was about to start. <i>Therefore her</i> palms are sweaty. | Laura got a lot of mail today. <i>Therefore her</i> palms are sweaty |
| | Laura got a lot of mail today. <i>Her</i> friends had remembered <i>her</i> birthday. | Mary's exam was about to start. <i>Her</i> friends had remembered <i>her</i> birthday. |
| Incohesive | Mary's exam was about to start. The palms are sweaty. | Laura got a lot of mail today. The palms are sweaty. |
| | Laura got a lot of mail today. Some friends had remembered the birthday. | Mary's exam was about to start. Some friends had remembered the birthday. |

As can be seen, lexical cohesion has opposing effects depending on pragmatic coherence: In the incoherent trials, cohesion leads to a pragmatic garden-path effect, whereas in coherent trials, cohesion facilitates inferencing (Ferstl et al., 2002). For instance, in the first example presented in Table 8 (coherent/cohesive), the connective '*therefore*' indicates the type of inference that needs to be made (i.e. causal), but not its pragmatic content. This content needs to be inferred by the reader on the basis of world knowledge (i.e. palms become sweaty when people are nervous; exams make people nervous). The same example that is missing a cohesive tie in Table 8 (coherent/incohesive) provides a case where this lexical device does not indicate the type of inference, but nonetheless, the same causal inference is made on the basis of world knowledge. Thus, when a cohesive tie is present in the coherent trial, the inference becomes easier because of the lexical information provided by the tie. However, when a cohesive tie erroneously indicates a relationship between two incoherent sentences (incoherent/cohesive), processing is made more difficult, perhaps due to a competition between pragmatic incoherence and linguistic cohesion (Ferstl & von Cramon, 2001).

This paradigm has been used in a number of neuropsychological and neuroimaging experiments (Ferstl & von Cramon, 2001; Ferstl, Guthke, & von Cramon, 2002). In the fMRI experiment (Ferstl & von Cramon, 2001) behavioural data confirmed that for the coherent sentence pairs, cohesion facilitated processing (as indexed by a speeding of reading times). For the incoherent sentence pairs, the cohesive ties produced a mismatch between lexical and pragmatic information, indexed by an increase in reading times and error rates for this condition.

The two studies described below sought to explore component processes that may be related to performance in the coherence and cohesion paradigm. Experiment 3 took measures of decoding skill, vocabulary, working memory and off-line comprehension skill. If sluggish or laboured decoding processes monopolise working memory resources and constrain higher-order comprehension processes such as sentence integration and inferencing (c.f. Perfetti & Hart, 2002; Perfetti, 2007), then decoding skill should be related to the size of the interaction between coherence and cohesion. If the relationship between working memory and reading comprehension skills, such as sentence integration and inferencing, is determined solely by language ability (MacDonald & Christiansen, 2002) and/or access to long term memory (Ericsson & Kintsch, 1995), then the magnitude of the interaction between coherence and cohesion should be predicted by decoding ability, vocabulary and verbal intelligence only.

Working memory was the only variable that was related to the size of the interaction between coherence and cohesion, providing support for the capacity constrained model of comprehension skill (Just & Carpenter, 1992). This finding motivated

Experiment 4, the aim of which was to explore the predictions of the long-term working memory model (Ericsson & Kintsch, 1995). The model proposes that the relationship between working memory and reading comprehension is driven by access to LTM. As such, this model would predict that semantic processing skills and semantic access would be related to the effect. The results of Experiment 4 supported this contention, and showed that scores on a verbal fluency task were able to predict variance in the magnitude of the interaction between coherence and cohesion.

5.3 Experiment 3: Coherence, Cohesion and Component Skills

5.3.1 Introduction

The aim of the present study was to assess the influences of reading comprehension skill, vocabulary, decoding ability and working memory on reading times and error rates in the coherence and cohesion paradigm. The hypotheses relating to each of these variables are stated below.

Reading Comprehension Skill

In a story comprehension task, Long and Chong (2001) found individual differences in the maintenance of local and global coherence in adult readers classified as good and poor comprehenders. Participants read stories in which the actions of a character were either consistent or inconsistent with earlier descriptions of them. The description and the action were either adjacent in the text (local coherence) or were separated by intervening text (global coherence). Only the good comprehenders were affected by incongruities at a

global level, whereas both good and poor comprehenders showed disrupted processing for local (sentential) inconsistencies.

The two-sentence stories in the present study are designed so that local and global coherence coincide. If the findings of Long and Chong are due to temporal distance effects, then comprehension skill should not influence the speed with which readers process the sentences in the four conditions, and it should not affect the accuracy of their coherence judgements. However, the stories in the present experiment require readers to make inferences on the basis of world knowledge. It is known that less-skilled readers make fewer of these inferences than do skilled readers (e.g. Long, Oppy, & Seely, 1997). Further, poor comprehenders have been found to have difficulties answering questions that require such an inference to be made, even when the text is available during questioning (Oakhill, 1984). The prediction, therefore, is that comprehension skill will influence performance in the present study by virtue of its reliance on knowledge based inferences, regardless of the fact that the sentences are adjacent rather than separated by intervening text.

Working memory

In an error detection paradigm, Oakhill, Hartt, & Samols (1995) showed that children found it more difficult to detect inconsistencies that were separated by several sentences than when they were presented in adjacent sentences. This distance effect was more pronounced in children with comprehension weaknesses than it was in children classified as good comprehenders. Oakhill (1996) concluded that less skilled

comprehenders do not have *general* deficits integrating incoming information with their mental model, but rather, that difficulties are revealed when memory demands are great.

As stated, the materials used in the present study were designed so that local and global coherence coincided in the same event – at the target sentence. The demands placed on working memory are relatively small in the current task, at least in comparison to longer, more naturalistic texts. Even so, the present task requires participants to read and store the context sentence at a level that is sufficient to allow an assessment of coherence between it and the target sentence.

Ferstl and von Cramon (2001) found evidence to suggest that brain regions that are implicated in a variety of tasks targeting executive functions are also recruited in the conflict condition of the coherence and cohesion paradigm. In particular, when sentences do not cohere yet are erroneously linked with a cohesive tie, prefrontal regions become preferentially activated. Since the sentence span task is thought to reflect aspects of executive functioning (Whitney, Arnett, Driver, & Budd, 2001), it is likely that performance on the sentence span task will relate to performance on the coherence judgement paradigm. Further, there is evidence that performance on the sentence span task reflects speed of processing. For instance, Whitney et al. (2001) found that a measure of speeded decision making contributed a small but significant amount of variance to performance on the sentence span task ($R^2 = .10$). The predictions with regard to working memory capacity in the present study are as follows: In trials that are coherent and incohesive (i.e. coherence is not made explicit by the presence of a cohesive tie), judgements of coherence need to be made solely on the basis of pragmatic information. In this condition,

those with a high working memory capacity will read the target sentence more quickly than those with a low working memory capacity as they should be able to draw on world knowledge more effectively and should be able ascertain the relationship between the two sentences more efficiently. In trials that are incoherent but are misleadingly linked with a cohesive tie (i.e. incoherent and cohesive), the cohesive tie should cause readers to reprocess the target sentence. Those with a high working memory capacity should be able to engage in this process more efficiently than those with a lower working memory capacity, presumably because the target sentence will be readily available to them for reprocessing.

The capacity constrained model of comprehension skill (Just & Carpenter, 1992) would predict a direct relationship between sentence span and performances in the coherence and cohesion paradigm. By attributing differences in working memory and reading comprehension to differences in language ability, the distributed learning account (MacDonald & Christiansen, 2002) would predict that performances in the coherence and cohesion paradigm would be mediated by language processing ability (i.e. decoding ability and/or vocabulary). The long-term working memory model (Ericsson & Kintsch, 1995) would predict that performances on the coherence and cohesion paradigm task should be mediated by access to world knowledge. In this experiment, access to world knowledge is rudimentarily indicated by performances in the vocabulary component of the Nelson Denny reading test.

Decoding skill and vocabulary

The ability to read words quickly and accurately is critical for successful reading comprehension. If lexical access does not occur fluently then resources needed for higher-order comprehension processes (such as inference generation) may be taken-up by lexical access and/or decoding (Perfetti & Hart, 2002; Perfetti, 2007). If this proposal is true, then performances on the vocabulary component of the Nelson Denny Reading Test, performance on the decoding tasks and performances on the sentence span task should each contribute unique variance to performance in the coherence and cohesion paradigm.

5.3.2 Method

Participants

Fifty two participants volunteered for the experiment. The 43 women and 9 men received either £10 or course credits for their participation. A total of 15 participants from the University of Sussex and 37 North American students from International Study Centre for Queens University were recruited for the experiment. Their ages ranged from 17 to 45 years of age ($M = 19.23$; $SD = 4.23$).

Materials and Procedure

The experiment took place over two sessions. The first session was run in groups, during which participants completed the 2 word-level tasks and the vocabulary and comprehension components of the Nelson-Denny reading test.

Session One (1 hour total)

Decoding Tasks

Decoding skill was measured using two tasks obtained from Charles Perfetti and Leslie Hart. These tasks were originally designed for use with American participants, and were adapted so that they could be used to assess speakers of British English. Both of these tasks require participants to detect phonologically legal yet orthographically incorrect words (i.e. pseudohomophones such as “EER”). They were administered exactly as described in Experiment 1.

Performance on the two pseudohomophone judgement tasks shared a significant correlation with one another ($r = .45, p < .001$). Scores on these tasks were transformed into z-scores, summed and divided by two, providing a composite measure of decoding skill (hereafter referred to as ‘decoding composite’).

Vocabulary and Comprehension

The Nelson-Denny Reading test (Brown et al., 1993) was used to measure comprehension, vocabulary and reading rate. Test characteristics and method of administration were exactly as described in Experiment 1.

Session 2 (45 minutes total)

Session Two was run individually on a participant-by-participant basis, usually within two days of participation in session one. The order in which participants completed the tasks is as presented below.

Working Memory

The sentence span task devised by Sarah McCallum (2002) was used to assess working memory capacity. Test characteristics and administration procedures are as described in Experiment 1.

Coherence & Cohesion Paradigm

As stated in the introduction, the task uses a 2 (coherence) x 2 (cohesion) within-subjects design. The paradigm was produced in four different versions (see appendices) so that the order of trial presentations was counterbalanced across participants. The E-prime stimulus presentation package was used to present the stimuli on a laptop computer. 'Yes' and 'no' keys were assigned to either the 'D' or 'L' buttons (counterbalanced across participants). The 'next' button was assigned to the space bar key.

Instructions were given both orally and in written form. Participants were instructed to read the first (context) sentence on the screen and to press 'next' as soon as they had done so. They were then instructed to read the second (target) sentence, and to press 'next' as soon as they had done so. The time between target onset and the button press was taken as the reading time measure. After the 2nd 'next' press, a question mark then appeared in the middle of the screen and participants were instructed to press 'yes' if they felt the sentences "made sense together" or 'no' if they felt they did not. Coherent

sentence pairs should always be answered ‘yes’, even in the absence of a cohesive tie. Incoherent sentence pairs should always be answered ‘no’, even when they are misleadingly linked with a cohesive tie.

Participants undertook eight practice trials to orient them to the demands of the task, after which they were asked if they had any questions. If the participant was comfortable with the procedure they then took the test proper. The task was split into four blocks, each comprised of 24 trials. Equal numbers of trials from the four conditions appeared in each block. Trials were presented in a pseudo-random order so that successive trials were not from the same condition and not more than three successive trials required the same answer.

5.3.3 Results

Means and standard deviations of performances on the individual difference tasks are presented in Table 9.

Table 9. Descriptive Statistics for Experiment 3 (N=52)

| | Mean | SD |
|---------------------------|--------|-------|
| ND Comprehension | 231.29 | 15.34 |
| ND Vocabulary | 237.08 | 13.80 |
| Working Memory | 4.58 | 0.95 |
| Transformation task | 0.08 | 0.04 |
| Pseudohomophone/3 | 46.54 | 6.80 |
| Pseudohomophone 3 minutes | 25.40 | 6.09 |

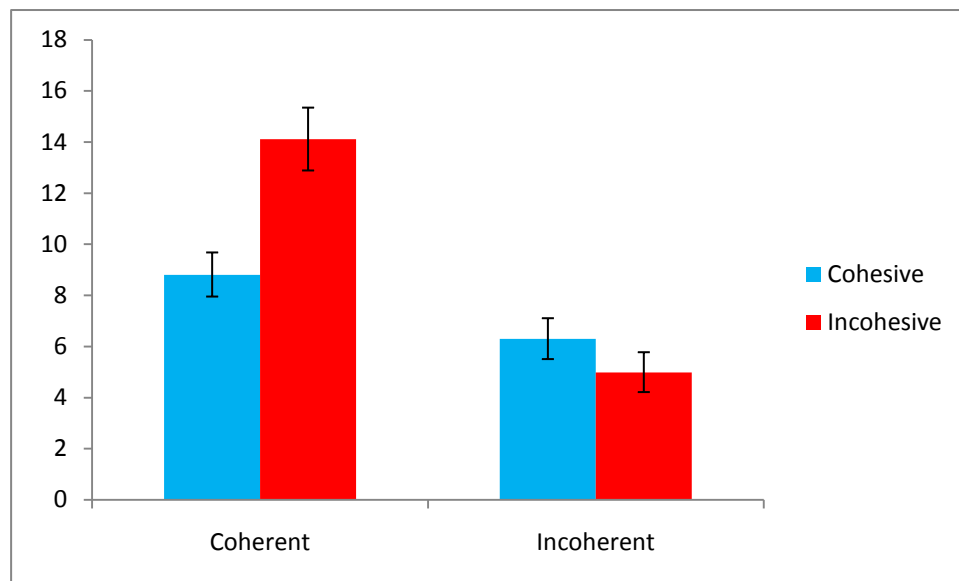
There were 3 trials that contained culturally specific terms that none of the Canadian participants (n = 37) understood (‘plasters’, ‘lorries’ and ‘midges’). As access to

the meaning of these words was deemed crucial for successful comprehension, these trials were excluded for these participants (111 data points in total) before error percentages and mean RTs were calculated. Error percentages and mean RTs are presented in Table 10.

Table 10. Accuracy and reading times for all participants in all conditions (mean(sd); n=52)

| | Coherent | | Incoherent | |
|--|---------------|--------------|--------------|---------------|
| | Cohesive | Incohesive | Cohesive | Incohesive |
| Percent error: | 8.81 (6.22) | 14.11 (8.89) | 6.30 (5.83) | 4.99 (5.59) |
| Target sentence reading times (ms): | 2134 (491) | 2014 (489) | 2221 (480) | 1930 (413) |
| Std. Residuals of target sentence reading times: | -.0401 (.231) | .0498(.329) | .1422 (.246) | -.0178 (.414) |

Figure 5. Mean percentage error rates for all conditions (n=52). Error bars represent one standard error above and below the mean



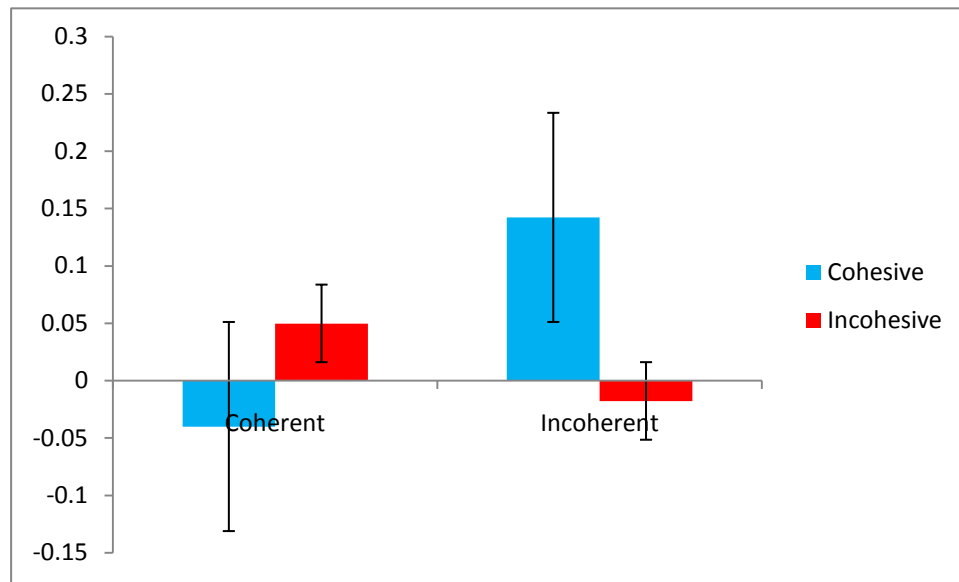
Overall, error rates were fairly low, with under 9% of trials answered incorrectly. The accuracy data were subjected to a 2 x 2 analysis of variance (ANOVA). There was a main effect of coherence ($F(1, 51) = 34.78, p < .001$). Participants were more accurate on incoherent trials than they were on coherent trials (see Figure 5). There was a main effect of cohesion ($F(1, 51) = 6.19, p < .02$), showing that participants' judgements of coherence were, overall, more accurate for sentence pairs that were bound with cohesive ties. There was also an interaction between coherence and cohesion ($F(1, 51) = 12.73, p = .001$). This interaction was driven by differences in the coherent condition, however. Thus, for sentence pairs that were coherent, cohesive ties served to make judgements of coherence more accurate, whereas their absence in these trials impaired participants' judgements, leading to an increase in errors ($t(52) = 3.76, p < .001$). When cohesive ties linked incoherent sentence pairs, these were erroneously judged to be coherent more often than when they were not linked with cohesive ties, though this effect was non-significant ($t(52) = 1.30, p = .20$).

For the analyses of reading times (RTs) incorrect trials were excluded (417/4881; 8.5%). Second, abnormally long trials were excluded (RTs above 5000 ms; 13/4464; 0.29%). Last, mean RTs to target sentences were calculated for each participant across conditions. Trials that were 2.5 standard deviations above their mean RT were replaced with a cut-off value (i.e. 2.5 SD above their mean reading time). In total, 4451 trials were used for the analysis of reading times, 116 of which had been replaced with each individual's 2.5sd cutoff (2.6% in total). Ninety-one percent of trials were retained for analysis.

Two reviewers noted that most of the sentence pairs in the coherent/incohesive conditions were perfectly acceptable, but some were quite obscure. In acceptable pairs, connectives in the target sentence were omitted, but their content made the intended connection clear (e.g. “It has been raining for days. The paths have turned to mud.”). In the anomalous pairs, the issue concerned replacing possessive pronouns (such as ‘her’) with the definite article, ‘the’. (e.g. “Marilyn has problems with her son. The teacher recommends a book about parenting.”). Such constructions rarely, if ever, appear in natural text or discourse, and would be judged as obscure by most readers. As a consequence, one cannot be sure whether or not differences in mean RT between conditions are driven by the experimental manipulations or by a few anomalous items. To address this issue, the RT data were analysed both by subject means and by items, with significant effects by items acting as a reliability check for any significant effects observed at the subject level.

As cohesive target sentences were longer than incohesive target sentences, it was necessary to factor-out sentence length for these analyses. For each participant separately, the reading times for target sentences were regressed on the number of characters (including spaces) in each sentence (i.e. sentence length). The standardised residuals that resulted from this procedure were then used for the following analyses. A graphical representation of the mean reading time residuals in the 4 conditions is presented in Figure 6.

Figure 6. Mean standardised residuals of target sentence reading times (n=52). Error bars represent one standard error above or below the mean



The mean standardised residuals for each participant were then used in a 2 (coherence) x 2 (cohesion) ANOVA. The interaction between coherence and cohesion depicted in Figure 6 was significant by subjects ($F_1(1, 51) = 18.07, p < .001$) and by items ($F_2(1, 1002) = 13.08, p < .001$). Thus, cohesive ties facilitated participants' comprehension of coherent sentence pairs, whereas their presence in incoherent sentence pairs impeded their comprehension. This impedance may stem from the fact that the coherence gap was contradicted by the presence of a coherence marker, causing readers to consider these cases in more depth and read them for relatively longer.

As stated in the introduction, it was hypothesised that readers' comprehension skill may affect judgements of coherence. To assess this, the sample of readers was divided into 'high' and 'low' comprehenders on the basis of their reading comprehension

scaled score falling above or below the group's median score for the comprehension component of the Nelson Denny (group median = 236.50).

A 2 (coherence) x 2 (cohesion) x 2 (ND comprehension group) mixed-design ANOVA, with ND comprehension group as the between subjects variable, was conducted on the standardised reading time residuals. There was no interaction between comprehension group and coherence by subjects ($F_1(1, 50) = 0.00, p = .99$), or by items ($F_2(1, 1001) = 2.48, p = .12$). There was no interaction between comprehension group and cohesion by subjects ($F_1(1, 50) = 2.56, p = .12$) or by items ($F_2(1, 1001) = 0.33, p = .57$). There was no 3-way interaction between coherence, cohesion and comprehension group by subjects ($F_1(1, 50) = 1.58, p = .22$) or by items ($F_2(1, 1001) = 0.30, p = .59$).

It was hypothesised that working memory capacity would play a role in participants' judgements of coherence. Based on performances on the sentence span task, the participants were placed into 'low' and 'high' working memory groups. Those who recalled either 3 or 4 items in the sentence span task were placed in the 'low working memory' group, those who recalled either 5 or 6 items were placed in the 'high working memory' group. The component reading characteristics of these two groups are presented in Table 11. Significant differences between the groups (as assessed by independent samples t-tests) are highlighted with asterisks.

Table 11. Component reading skills of working memory groups (mean(sd))

| | Working Memory Group | |
|-------------------------------------|----------------------|------------------|
| | High (n =28) | Low (n = 24) |
| Mean ND comprehension scaled score: | 237.43 (9.65)** | 224.13 (17.70)** |
| Mean ND vocabulary scaled score: | 240.68 (11.19)* | 232.88 (15.52)* |
| Mean working memory (max=6): | 5.29 (0.46)*** | 3.67 (0.48)*** |
| Mean decoding composite (Z-scores): | .2483 (0.56)* | -.2897 (1.04)* |

*** $p < .001$; ** $p < .01$; * $p < .05$.

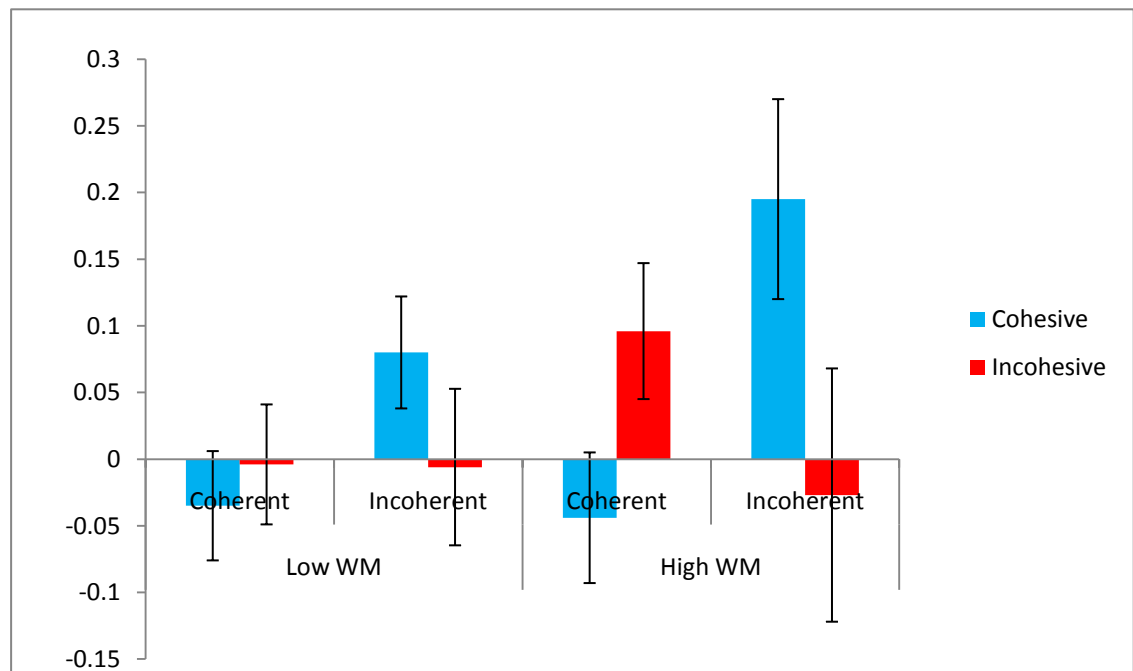
As shown in Table 11, the high working memory group performed significantly better than the low working memory group on the comprehension component of the Nelson Denny Reading Test ($t(50) = 3.43, p < .01$); on the vocabulary component of the Nelson Denny Reading Test ($t(50) = 2.10, p < .05$); on the working memory task ($t(50) = 12.38, p < .001$); and on the decoding composite ($t(50) = 2.37, p < .05$).

A 2 (coherence) x 2 (cohesion) x 2 (working memory group) mixed-design ANOVA, with working memory group as the between subjects factor, was conducted on the length-corrected residuals of the reading time data. There was no interaction between coherence and working memory group by subjects ($F_1(1, 50) = 0.00, p = .99$) or by items ($F_2(1, 1001) = 0.00, p = 1$), and there was no interaction between cohesion and working memory group by subjects ($F_1(1, 50) = 0.04, p = .84$) or by items ($F_2(1, 1001) = 0.74, p = .39$). This analysis did reveal, however, a significant 3-way interaction between coherence, cohesion and working memory group by subjects ($F_1(1, 50) = 4.68, p < .05$). This 3-way interaction approached significance when analysed by items ($F_2(1, 1001) = 3.19, p = .07$).

These findings suggested that the two groups responded differently to the experimental manipulations.

A 2 (coherence) x 2 (cohesion) ANOVA was conducted on the standardised reading time residuals of the two groups separately. Neither group showed main effects of coherence or cohesion (all $F_s < 1.00$). The low working memory group did not show an interaction between coherence and cohesion by subjects ($F_1(1, 23) = 1.47, p = .24$), or by items ($F_2(1, 458) = 1.10, p = .27$). In contrast, the high working memory group showed a strong interaction between coherence and cohesion by subjects ($F_1(1, 27) = 31.16, p < .001$), and by items ($F_2(1, 543) = 15.50, p < .001$). See Figure 7 for a graphical representation of these findings.

Figure 7. Mean residuals of target sentence reading times for the high working memory group (n=28) and the low working memory group (n=24). Error bars represent one standard error above and below the mean



As can be seen in Figure 7, when incoherent sentence pairs were (erroneously) linked with cohesive ties, the high working memory group read the target sentence relatively more slowly than they did in trials where they were not linked with cohesive ties. This suggests that the high working memory group could not help but be influenced by the pragmatic function that these devices serve (i.e. to link adjacent clauses). Conversely, the low working memory group did not show the same effect, at least not to the same extent as the high working memory group.

Cohesive ties act as a prompt for readers to link two adjacent clauses *explicitly*. In trials where this device facilitates comprehension (i.e. by linking sentences that cohere), high working memory readers are able to assess coherence more quickly (as indexed by a relative speeding of reading times). In trials where this device acts to impede comprehension (i.e. linking sentences that do not cohere), high working memory readers demonstrate this impedance by a relative *slowing* of their reading of the incoherent target sentence. Though the low working memory group show similar patterns, these are not pronounced enough to produce statistically significant differences between conditions.

To assess the contribution of the componential variables to the size of the observed interaction between coherence and cohesion, a step-wise hierarchical multiple regression analysis was conducted using the interaction value of the standardised residuals as the dependent variable and working memory, comprehension, vocabulary and word-level tasks as predictors.

The interaction value was ascertained for each participant using the residuals of target sentence reading times in each of the four conditions. The equation for this is as follows:

$$\text{Absolute interaction value} = (\text{Incoherent Cohesive residuals} - \text{Incoherent Incohesive residuals}) - (\text{Coherent Cohesive residuals} - \text{Coherent Incohesive residuals}).$$

A correlation matrix depicting the relationships between this variable and the component measures is presented in Table 12.

Table 12. Pearson's *r* correlation coefficients between component variables and the interaction values (*n*=52)

| | ND Comprehension | ND Vocabulary | Working Memory | Decoding Composite | Interaction Value |
|--------------------|------------------|---------------|----------------|--------------------|-------------------|
| ND Comprehension | 1 | | | | |
| ND Vocabulary | .66** | 1 | | | |
| Working Memory | .41** | .31* | 1 | | |
| Decoding Composite | .48** | .51** | .40** | 1 | |
| Interaction Value | .20 | .06 | .33* | .30* | 1 |

***p*<.01; **p*<.05.

The interaction value was used as the dependent variable in a stepwise hierarchical multiple regression analysis. The component measures that shared significant correlations with the interaction value were used as predictors in the regression model,

with their order of entry determined by the size of the correlation coefficients they shared with the DV. This analysis is presented in Table 13.

Table 13. Stepwise hierarchical regression analysis with absolute interaction value as the dependent variable (n=52)

| | <i>B</i> | <i>SE B</i> | β | ΔR^2 |
|--------------------|----------|-------------|--------------------|--------------------|
| Step 1 | | | | |
| Constant | -.424 | .279 | | |
| Working memory | .148 | .060 | .329* | .108* |
| Step 2 | | | | |
| Constant | -.259 | .300 | | |
| Working memory | .112 | .065 | .248^ | |
| Decoding composite | .101 | .072 | .203 ^{ns} | .035 ^{ns} |

* $p < .02$; ^ $p < .10$.

As can be seen in Table 13, working memory accounts for just under 11% of the variance in interaction effect size. In the second model, the decoding composite did not predict significant additional variance in the size of the interaction.

5.3.4 Conclusions

As shown in Figures 5 and 6, there was a significant 2-way interaction between coherence and cohesion in both the error data and standardised reading time data, replicating the findings of Ferstl and von Cramon (2001). For the error data, this effect was driven by differences in error rates in the coherent conditions, with participants making more errors when coherent sentences were not linked with cohesive ties in

comparison to sentences in which coherence was reinforced with a tie. Although participants made more errors when incoherent sentences were misleadingly linked with cohesive ties than when they were not, this effect was minimal and did not reach statistical significance.

For the standardised reading time data, significant differences were found in both the coherent and incoherent conditions. First, sentences that were coherent but were not linked with cohesive ties were read relatively more slowly than were sentences in which coherence was reinforced with a cohesive tie. Thus, lexical cohesion reduces the costs associated with integrating pragmatically coherent sentences. Second, sentences that did not cohere but were erroneously linked with cohesive ties were read relatively more slowly than incoherent sentences that were not linked with cohesive ties. The inclusion of cohesive ties in this condition resulted in a conflict between pragmatic and lexical information, increasing processing times.

The 3-way interaction between coherence, cohesion and reading comprehension group hypothesised in the introduction was not obtained. However, there was a 3-way interaction between coherence, cohesion and working memory group in the length-corrected, standardised reading time data, suggesting that the two working memory groups performed quite differently on the task.

Cohesive ties act as an explicit prompt for readers to link two adjacent clauses. Both groups of readers read the sentences that cohered, and in which coherence was reinforced with a cohesive tie, relatively more quickly. Sentences that did not cohere and in which their incoherence was reinforced by the *absence* of a cohesive tie were read

relatively more quickly by those with a high working memory span than those with a low working memory span. When reading times were transformed into standardised values corrected for sentence length, those with higher working memory spans were able to establish the relation between the two sentences more efficiently than were those with lower working memory spans.

More interesting and unexpected effects were revealed in conditions that do not adhere to normal or 'standard' sentence constructions. Sentences that cohere but in which coherence is not reinforced by a cohesive tie were read relatively more *slowly* by participants with higher working memory spans than those with lower working memory spans. Similarly, sentences that did not cohere yet were erroneously linked with cohesive ties, making their incoherence ambiguous, were also read relatively more slowly by participants with higher working memory spans.

In support of these findings, the stepwise multiple regression analysis presented in Table 13 showed that the only significant predictor of the interaction effect between coherence and cohesion in the standardised reading time values was working memory span. As stated in the introduction, performance on the sentence span task is thought to reflect functional aspects of the central executive (Whitney et al., 2001), a module that is believed to be involved in comprehension monitoring (Sesma, Mahone, Levine, Eason, & Cutting, 2009). Comprehension monitoring is a metacognitive process that directs a reader's cognitive resources as they attend to incoming information. This process appears to involve some kind of triggering mechanism, or a conscious recognition by the reader of a failure to understand the text message (Wagoner, 1983). In the present case, this trigger

appears to be the presence of a cohesive tie that falsely indicates coherence between two sentences, and the conspicuous absence of a tie in cases where pragmatic inferences indicate that two sentences do cohere. The present data suggest that readers' sensitivity to these triggers is dependent upon working memory capacity.

Those with a high working memory capacity are not only able to comprehend and link the sentences with relative ease in the 'usual' conditions (i.e. where coherence is reinforced with a cohesive tie and incoherence is reinforced by the absence of a tie), but also that they are able to monitor their comprehension much more effectively in the 'unusual' conditions. In cases where a cohesive tie serves to erroneously link adjacent, incoherent sentences, the erroneous tie triggers a monitoring warning in the high-span readers, causing them to evaluate the second sentence in more depth. Similarly, in cases where the sentences cohere but are not linked with a cohesive tie, the conspicuous absence of the tie triggers a monitoring warning, causing a relative slowing of reading times for the target sentence. The same cannot be said of the low-span readers who were, in general, less responsive to the differences between the conditions.

The findings of experiment are best explained by the capacity constrained model of reading comprehension (Just & Carpenter, 1992), which predicted a direct relationship between sentence span performances and performances on the coherence and cohesion paradigm. The fact that vocabulary and word-level abilities were not related to variance in the size of the interaction between coherence and cohesion does not lend support to the distributed learning model (MacDonald & Christiansen, 2002), the long-term working

memory model (Ericsson & Kintsch, 1995), or the Lexical Quality hypothesis (Perfetti & Hart, 2002; Perfetti, 2007).

5.4 Experiment 4: Coherence, Cohesion, Component Skills, Semantic Access and Verbal Intelligence

5.4.1 Introduction

Experiment Three demonstrated that when reading times for the four conditions were transformed into standardised values, those with a high working memory capacity showed a larger interaction between coherence and cohesion than did those with a low working memory capacity. It was argued that when participants encountered atypical sentence pairs in the paradigm (i.e. sentence pairs that cohere but are missing cohesive ties, and sentences that are incoherent but are erroneously linked with cohesive ties) the high-working memory participants engaged in some form of monitoring process as a consequence of this atypicality. This was evidenced by the fact that the high working memory group showed a much larger interaction effect than the low-working memory group, suggesting that they were taking relatively longer to process target sentences that appeared in the atypical conditions.

Implicit in this contention is that high capacity readers are elaborating on the pragmatic relationships between the first and second sentences to a greater degree than are low capacity readers. Assessments of pragmatic coherence for sentence pairs that are

coherent but which are not linked with a cohesive tie require readers to reflect on world knowledge rather than lexical information. It may be that those with a high working memory capacity are able to engage in a greater degree of semantic elaboration; linking incoming information to general or world knowledge. Under this proposal, the interaction between coherence and cohesion may be better described as an 'elaboration effect' rather than a 'monitoring effect'. The fact that working memory capacity does not interact with error rates in the coherence and cohesion paradigm accords with this proposal; the elaboration is not essential to coherence judgements. Rather, the larger effect sizes shown by the high-capacity subjects may be indicative of more methodical or exhaustive sentence integration operations.

The present experiment sought to determine whether the interaction between coherence and cohesion is driven by efficient semantic processing abilities rather than working memory capacity per se. There is evidence to suggest that less skilled comprehenders have deficits in semantic processing, and are less able to access knowledge of word meanings. Nation and Snowling (1999) compared children who differed with regard to comprehension skill but were matched for nonword decoding on a semantic priming task. The two groups showed comparable priming for items related by function (e.g. BROOM-FLOOR), but children with comprehension deficits did not show priming for category coordinates (e.g. AEROPLANE-TRAIN) *unless* the pairs shared high association strength (e.g. CAT-DOG). The authors concluded that in the absence of explicit co-occurrence, poor comprehenders are less sensitive to abstract semantic relations than are normal readers (Nation & Snowling, 1999).

Landi and Perfetti (2007) have tested this semantic deficit hypothesis in adult readers using behavioural and event related potential (ERP) techniques. The less skilled comprehenders in their study were readers with relatively high lexical skills (as assessed by decoding, spelling and vocabulary measures) but with poor comprehension skill (as assessed by the Nelson Denny reading test). As in Nation and Snowling's (1999) study, these participants made meaning decisions on word pairs (i.e. to state whether or not they were related) that were either categorically and associatively related or were related by category only. Based on the findings of Nation and Snowling (1999), the authors predicted that meaning decisions for word pairs that were both categorically and associatively related would be comparable across the two groups (as associative relations could facilitate the detection of semantic relatedness). However, they predicted that the less skilled comprehenders would perform more poorly when they had to base their decisions on categorical relations alone, as these decisions would necessitate a greater degree of semantic processing (Perfetti, 2007). The two groups did not differ in the accuracy of their judgements (both groups performed very well), but the skilled comprehenders performed the tasks more quickly. For both groups, correct decisions for pairs related by category alone were slower than decisions based on categorical and associative relations, but these decisions took, on average, 68ms longer for the skilled group, and 139ms longer for the less skilled group. Although the effect occurred in the predicted direction, it was subtle and failed to reach statistical significance (Perfetti, 2007).

More striking differences were observed in event-related brain potentials (ERPs) that were produced in response to the second words of each pair. ERPs are time-locked components of an electroencephalographic (EEG) signal produced by the brain and measured at the scalp. In language processing tasks, a negative-going deflection that peaks at around 400ms (the N400) has been shown to vary systematically with the processing of semantic information (e.g. Kutas & Hillyard, 1980). In priming studies, prior occurrence of an associatively related item (e.g. 'bee'), or a semantically related item (e.g. 'sugar') reduces the N400 amplitude to a given word (e.g. 'honey') (Kutas & Federmeier, 2000). In Landi and Perfetti's (2007) study, both groups of readers showed a reduction in N400 amplitude to the second word of each pair. This was reliable in both groups for word pairs that were related categorically (although the effect was larger for the high-skill group). However, the high-skill group, but not the low-skill group, showed an *additional* reduction in N400 amplitude to word pairs that were related both by association and category. This finding suggests that semantic processing involved in the appraisal of categorical relations were comparable in the two groups, but ERPs to associative relations are not. Thus, skilled comprehenders had a stronger meaning congruence response when the words were categorically related, and this effect was even stronger when there was an additional associative relation. In contrast, the less skilled group showed a congruence response to categorically related words, but this amplitude reduction was less pronounced than the skilled group's, and was not strengthened by additional associative relations (Perfetti, 2007). Taken together, the results of Nation and Snowling (1999) and

Landi and Perfetti (2007) suggest that semantic deficits, though subtle, are able to differentiate skilled and less skilled comprehenders, regardless of age.

There is evidence to suggest that semantic processes come into play when readers are comprehending texts that differ in coherence, and that background knowledge becomes especially important when reading texts judged to be of low coherence. McNamara and colleagues have manipulated text coherence by adding connectives, replacing pronouns with noun phrases and replacing words to increase coherence (McNamara & Kintsch, 1996). When modified biology texts were presented to eight grade readers, those with a low knowledge base benefited most from these manipulations, showing improvements in off-line comprehension measures (memory for the text and responses to multiple choice questions). However, readers with a high knowledge base demonstrated improved comprehension for texts that were of *low* coherence. These high knowledge readers demonstrated an improved conceptual understanding of the texts by performing better on questions related to bridging inferences, problem solving questions and accuracy in a keyword sorting task (where words that occurred frequently in the text had to be organised into groups on the basis of their relation to one another)(McNamara & Kintsch, 1996). McNamara (2001) attributed this low-coherence advantage for high knowledge readers to their use of knowledge to fill-in conceptual gaps in the texts: By interpolating background knowledge into the text representation, the high knowledge readers were able to create more elaborate and coherent discourse-level representations.

In the present case, the increase in the effect size shown by high capacity readers may be indicative of a greater degree of elaboration; linking the text propositions to prior

knowledge through semantic associations. That is, it may be that these readers are able to create a more elaborate mental representation of the short passages, and that this ability is governed by semantic processing skills. These skills come into play in the atypical conditions; sentences in which congruence needs to be inferred on the basis of pragmatic information in the absence of lexical information (the coherent and incohesive condition), and sentences in which coherence needs to be rejected when a lexical tie falsely indicates a pragmatic connection (the incoherent and cohesive condition). Specifically, the proposal is that the absence of a cohesive tie leads the reader to reflect on background knowledge to make the appropriate connection, and the presence of an erroneous cohesive tie triggers an appraisal made on the basis of background knowledge. If so, then measures of semantic ability should predict unique variance in the size of the interaction effect, over and above that predicted by working memory capacity.

To explore this possibility, the paradigm was re-run with new participants and additional measures. Two of these were designed to measure semantic access through the speed and accuracy with which participants were able to make synonym and hypernym judgements. These tasks were chosen because readers often need to represent noun phrases at a higher-level in order to make anaphoric links between sentences, and there is evidence to suggest that readers with comprehension difficulties do not make these links as efficiently as skilled comprehenders. For example, Perfetti, Yang, & Schmalhofer (2008) have found that skilled readers show both N400 and P300 responses to words that are referents of previously presented information (e.g. “Allen... rushed off to the emergency room. The *hospital* had a long waiting line”). The P300 component has been associated

with memory updating, and is believed to reflect recognition processes (e.g. Miller, Baratta, Wynveen, & Rosenfeld, 2001). Perfetti et al. (2008) proposed that the P300 responses were elicited by the word 'hospital' because it was anaphorically linked to the noun phrase 'emergency room', and that the N400 effect represented the processing costs associated with integrating those two meanings into a discourse-level representation (Perfetti et al., 2008). Less skilled readers did show similar electrophysiological responses, but these were observed to be much reduced and temporally delayed in comparison to those produced by skilled readers. By requiring participants to represent given words at a higher-level, the synonym and hypernym judgement tasks used in the present study may provide a window into these processing differences. The prediction is that better performance (in terms of accuracy, speed, or both) on these tasks should be related to the size of the interaction between coherence and cohesion.

A third task sought to measure semantic access through the accuracy and speed with which participants are able to form meaningful categories by selecting the 'odd-one-out' from groups of semantically related words (e.g. flea, ball, kangaroo, and *car*). Again, if the interaction effect seen in the coherence and cohesion paradigm is attributable to efficient semantic processing and the ability to represent categorical relations at a higher-level, then performance on this task should predict unique variance in the size of the interaction effect, with higher performance predicting a larger interaction effect.

A fourth task sought to examine semantic fluency through a 'free association' procedure, whereby participants were asked to produce semantically related words in

response to stimulus items taken from the Deese (1959), Roediger and McDermott (1995) paradigm. This task taps participants' ability to represent given items at a higher level by asking them to produce words that are related in meaning to the stimulus words. Better performance on this task should predict an increase in the interaction between coherence and cohesion, and should predict unique variance over and above that predicted by working memory capacity.

These tasks were intended to provide broad, exploratory assessments of semantic processing. These tasks tap semantic processes at lower- and higher-levels (e.g. lexical associations in the synonym task vs. pragmatic associations in the odd-one-out task and the free association task). Based on the assertion of the long-term working memory model (Ericsson & Kintsch, 1995) and the CI model (Kintsch, 1988; 1999) that inference generation is dependent upon access to long term memory and world knowledge, and previous demonstrations of the relation between semantic ability (Perfetti et al., 2008; Nation and Snowling, 1999), access to knowledge (McNamara, 2001), and reading comprehension ability, it was predicted that performances on one, or all, of these tasks would be related to the size of the interaction between coherence and cohesion. Predictions about how each of these tasks may independently relate to this interaction effect were not made a-priori, however.

To test whether variance in the interaction between coherence and cohesion could be attributed to verbal intelligence rather than aspects of semantic ability (as would be predicted by the distributed learning model of MacDonald and Christiansen (2002)), the National Adult Reading Test (Nelson, 1991) was administered to provide an assessment of

verbal intelligence. This measure was included along with the other measures as part of a large, correlational analysis (see below).

5.4.2 Method

Participants

Sixty three adults volunteered for the experiment. The 41 women and 22 men received £10 for their participation. The participants' ages ranged from 17 years to 41 years ($M = 25$, $SD = 4.78$). Twenty four of the participants were North American exchange students on summer school programs of various disciplines. The remaining 39 were a mixture of undergraduate and post-graduate students, clerical employees at The University of Sussex and members of the general public.

Materials and Procedure

The experiment took place over two sessions. The first session was run in groups, during which participants completed the two decoding tasks and the vocabulary and comprehension components of the Nelson-Denny reading test.

Session One (one hour total)

Decoding tasks

The two decoding tasks described in Experiments 1, 2 and 3 were also used in the present experiment. They were administered exactly as described in Experiment 1.

Vocabulary and Comprehension

Participants then took the vocabulary component of the Nelson Denny Reading Test (Brown, Fischco & Hanna, 1993), followed immediately by the comprehension component. They were administered exactly as described in Experiment 1.

Session Two (one hour total)

Session two was run individually on a participant-by-participant basis, usually within two days of participation in session one. The order in which participants completed the tasks is as presented below.

Working Memory

The sentence span task devised by Sarah McCallum (2002) that was used to assess working memory capacity in Experiments 1, 2 and 3 was also used in the present experiment. The task was administered as described in Experiment 1.

Odd-one-out Task

Participants were presented with 40 groups of 4-word rows. One of the words did not fit with the other 3 when they were conceptualised as a category (e.g. “car, flea, ball, kangaroo”. Answer: car). Participants were required to circle the ‘odd-one-out’ in each of the rows, and were instructed to work as quickly as possible with a maximum of 5 minutes to complete the task. If the task was completed in less than 5 minutes, the participants were asked to let the experimenter know immediately. A total of 13 participants needed the entire 5 minutes to complete the task. Performance times for the other participants (n= 50) were recorded by the experimenter. Both accuracy and RTs were used in the analyses presented below.

Verbal Intelligence

Participants then took the National Adult Reading Test (NART) (Nelson, 1991). This test requires participants to read aloud a total of 50 irregular and infrequent words (e.g. 'Quadruped') and is used as a measure of verbal intelligence. Test characteristics and method of administration are exactly as described in Experiment 1.

Free-Association Task

Six words that were classified as 'critical associates' in the semantically related word-lists used in the Deese (1959); Roediger and McDermott (1995) paradigm were presented on separate pieces of paper ('smell', 'sweet', 'window', 'sleep', 'chair' and 'doctor'). Participants were instructed to produce as many words as possible that were related in meaning to the stimulus item, and were given one minute in which to do so. After one minute had passed they were instructed to turn the page to the next item and to repeat the procedure. A total of six words were used as stimulus items, meaning that the test took just over 6 minutes to complete.

Participants received one point for each word produced (not including repetitions). Affixes that were added to words that had already been produced were scored in two ways: Affixes added to words inflectionally (modifying tense, plurality etc.) were classified as repetitions and scored as zero. Affixes that were appended derivationally (modifying word-class, e.g. 'Flavour' – 'Flavoursome') were classified as new words and were awarded one point. None of the participants produced an unrelated word in any of the testing sessions.

Speeded Response Tasks

Participants were seated in front of a laptop informed that they would be taking the computer-based tasks. Response buttons for 'yes' and 'no' were assigned to either the 'L' or 'D' keys (counterbalanced across participants). A 'next' response label was assigned to the space-bar key. Participants were instructed that both RTs and responses were to be recorded in all of the subsequent tasks, and that they should therefore place their hands in a 'strategic position' so that they were able to respond as quickly as possible.

Synonym Judgement

Twenty trials of word-pairs that were either synonyms of each other or not (10 trials of synonyms and 10 trials of non-synonyms) were presented in a random order using E-Prime stimulus presentation software. Sentences such as, "Slim means thin" were presented on a computer screen, and participants were required to accept or reject them (yes or no button responses). Instructions were given both orally and in written form. A total of 4 practice trials were administered to orient the participants to the requirements of the task. Accuracy and RTs were recorded throughout. Incorrect trials were excluded, as were trials that fell 2.5 standard deviations above each participant's mean RT. Accuracy and mean RTs for correct trials were used in the analyses presented below.

Hypernym Judgement

Twenty trials of hypernym (10) or non-hypernym statements (10) were presented to participants in random order using E-Prime stimulus software package. Statements such as, "Breakfast is a type of meal," were accepted or rejected by yes/no button

presses. Instructions were given both orally and in written form. A total of 4 practice trials were administered to orient the participants to the requirements of the task. Accuracy and RTs were recorded throughout. Incorrect trials were excluded, as were trials that fell 2.5 standard deviations above each participant's mean RT. Mean RTs for correct trials were used in the analyses presented below.

Same/Different Judgement Control Task

E-Prime stimulus software package was used to present participants with 20 trials of word-pairs that were either identical or non-identical, in a random order. Participants were required to accept or reject statements such as, "Tree = Tree". Instructions were given both orally and in written form. Accuracy and RTs were recorded throughout. Incorrect trials were excluded, as were trials that fell 2.5 standard deviations above each participant's mean RT. Mean RTs for correct trials were used in the analyses presented below.

Coherence & Cohesion Paradigm

Participants then took part in the sentence comprehension task as described in Experiment 3.

5.4.3 Results

Means and standard deviations of performances on the individual difference tasks are presented in Table 14.

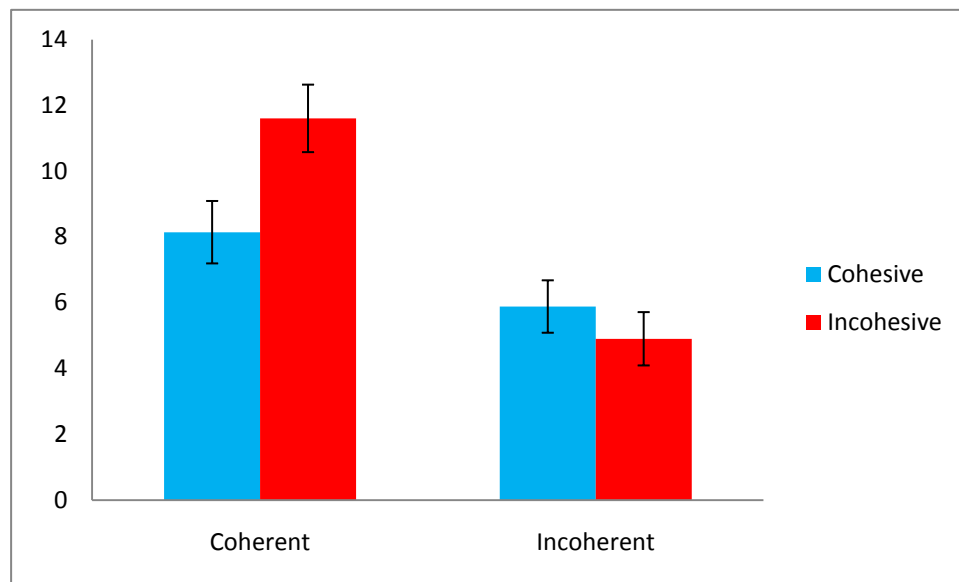
Table 14. Descriptive statistics for the component measures of Experiment 4

| | Mean | SD |
|---------------------------|---------|--------|
| ND Comprehension | 239.02 | 12.51 |
| ND Vocabulary | 247.16 | 7.07 |
| NART | 31.60 | 9.99 |
| Working Memory | 4.41 | 1.01 |
| Pseudohomophone/3 | 50.11 | 2.23 |
| Pseudohomophone 3 minutes | 29.32 | 4.68 |
| Odd-one-out accuracy | 34.38 | 3.68 |
| Odd-one-out RT | 247.49 | 41.46 |
| Free association | 81.74 | 22.68 |
| Synonym accuracy | 18.32 | 1.09 |
| Synonym RT | 1411.41 | 339.30 |
| Hypernym accuracy | 19.49 | 0.72 |
| Hypernym RT | 1244.61 | 253.99 |
| Same/Different accuracy | 19.79 | 0.48 |
| Same/Different RT | 660.70 | 120.40 |

Three of the 96 sentence pairs were deemed to be culturally biased (containing the words ‘plasters’, ‘lorries’ and ‘midges’). As access to the meanings of these words was deemed crucial for comprehension, these trials were removed from the North American participants’ data prior to error rate and reading time calculations (69 data points in total). Error percentages and mean RTs are presented in Table 15.

Table 15. Accuracy and reading times for all participants in all conditions (mean(sd); n=63)

| | Coherent | | Incoherent | |
|-------------------------------------|---------------|-------------|-------------|--------------|
| | Cohesive | Incohesive | Cohesive | Incohesive |
| Percent error: | 8.14(7.53) | 11.60(8.15) | 5.88 (6.33) | 4.90(6.43) |
| Target sentence reading times (ms): | 1857(494) | 1725(450) | 1918(499) | 1703(484) |
| Std. Residuals of target sentence | | | | |
| RTs: | -.0304(.2171) | .0174(.217) | .0684(.208) | -.0427(.226) |

Figure 8. Mean percentage error rates for all conditions (n=63). Error bars indicate one standard error above and below the mean

Overall, the participants made relatively few errors, with under 8% of trials answered incorrectly (455/5979). The accuracy data were subjected to a 2 (coherent; incoherent) x 2 (cohesive; incohesive) analysis of variance (ANOVA). This revealed a main effect of coherence, ($F(1, 62) = 16.62, p < .001$), because participants were more accurate in incoherent trials than in coherent trials (see Figure 8). There was a main effect of

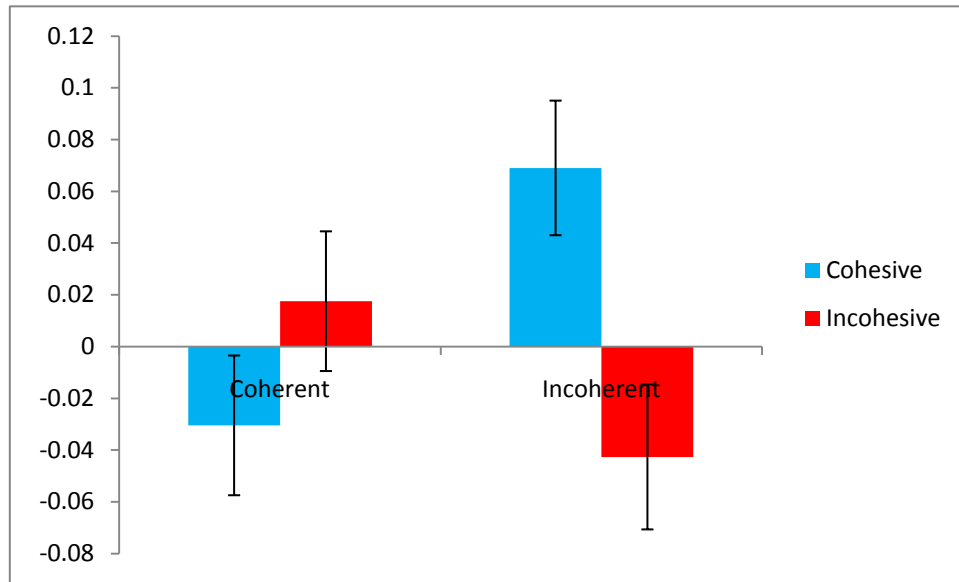
cohesion ($F(1, 62) = 4.05, p < .05$), because participants' judgements of coherence were, overall, more accurate for sentence pairs that were bound with cohesive ties. The ANOVA also revealed a significant interaction between coherence and cohesion ($F(1, 62) = 14.49, p < .001$). For sentence pairs that were coherent, cohesive ties served to make judgements of coherence more accurate, whereas their absence in these trials impaired participant's judgements of coherence. When cohesive ties linked incoherent sentence pairs, these were erroneously judged to be coherent more often than when they were not.

For the analyses of reading times, incorrect trials were excluded (455/5979; 7.6%). Mean reading times to target sentences were then calculated for each participant across conditions. Trials that fell 2.5 standard deviations above their mean reading time were replaced with that value (i.e. 2.5sd above the mean). In total, 5524 (92%) trials were used for the reading times analyses, 161 (2.91%) of which had been replaced with each individual's 2.5sd cut-off.

As in Experiment 3, the RT data were analysed by subject and by item. This was to ensure that any significant effects at the subject-level were driven by the experimental manipulations rather than anomalous sentence pairs at the item-level.

As cohesive target sentences were longer than incohesive target sentences, it was necessary to factor-out sentence length for these analyses. For each participant separately, the reading times for target sentences were regressed on the number of characters (including spaces) in each sentence (i.e. sentence length). The standardised residuals that resulted from this procedure were then used for the following analyses.

Figure 9. Mean standardised residuals of target sentence reading times (n=63). Error bars indicate one standard error above and below the mean



The mean residuals were used in a 2 (coherent; incoherent) x 2 (cohesive; incohesive) ANOVA. The interaction depicted in Figure 9 was found to be statistically significant by subjects ($F_1(1, 62) = 10.37, p = .002$), and by items ($F_2(1, 1272) = 82.92, p < .001$). Thus, cohesive ties facilitated participants' comprehension of coherent sentence pairs, whereas their presence in incoherent sentence pairs impeded their comprehension. This impedance may stem from the fact that the coherence gap was contradicted by the presence of a coherence marker, causing readers to consider these cases in more depth and read them for relatively longer.

On the basis of the findings in Experiment 3, it was predicted that working memory span would affect participants' judgements of coherence. Experiment 3 suggested that those with a high working memory capacity are more susceptible to the

manipulations of coherence and cohesion, possibly as a result of unhindered monitoring processes that are activated by these manipulations. As in the previous experiment, participants were divided into 'low' and 'high' working memory groups on the basis of performance in the sentence span task. Participants who recalled 3 or 4 items were placed in the low group ($n = 33$), and those who recalled 5 or 6 items were placed in the high group ($n = 30$). A 2 (coherent; incoherent) \times 2 (cohesive; incohesive) \times 2 (low working memory; high working memory) between subjects ANOVA was conducted on the standardised residuals. There was no interaction between coherence and working memory group by subjects ($F_1(1, 61) = 0.44, p = .51$) or by items ($F_2(1, 1271) = 1.47, p = .23$). There was no interaction between working memory group and cohesion by subjects ($F_1(1, 61) = 0.01, p = .91$), or by items ($F_2(1, 1271) = 0.00, p = .97$). Contrary to predictions, the ANOVA did not reveal a 3-way interaction between coherence, cohesion and working memory group by subjects ($F_1(1, 61) = 0.24, p = .63$), or by items ($F_2(1, 1271) = 1.47, p = .23$).

As stated, one of the motivations for administering the coherence and cohesion paradigm in this study was to explore the possibility that semantic access may play a role in the 3-way interaction observed in Experiment 3. As such, a number of tasks designed to assess lexical quality were conducted immediately before participants undertook the coherence and cohesion paradigm. To elucidate the relationships between these variables, a correlation matrix was produced (see Fig. 3). It should be noted that due to experimenter error, 4 participants were not graded on the Free Association task ($n = 58$).

In order to assess the relationships between each of the measures to variance in the CoCo interaction, it was first necessary to calculate the absolute value of the interaction effect for each participant. The equation for this is as follows:

Absolute interaction value = (Incoherent Cohesive residuals – Incoherent Incohesive residuals) – (Coherent Cohesive residuals – Coherent Incohesive residuals).

Table 16. Pearson's r correlation coefficients

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|----------------------------|--------|--------|--------|--------|--------|-------|--------|------|--------|------|--------|--------|---------|-------|------|------|----|
| 1. Vocab. Scaled | 1 | | | | | | | | | | | | | | | | |
| 2. Comp. scaled | .23^ | 1 | | | | | | | | | | | | | | | |
| 3. NART | .59*** | .22^ | 1 | | | | | | | | | | | | | | |
| 4. Working Memory | .26* | -.13 | .20 | 1 | | | | | | | | | | | | | |
| 5. Decoding Composite | .43*** | .16 | .47*** | .15 | 1 | | | | | | | | | | | | |
| 6. Same Different Accuracy | .03 | -.06 | .00 | .11 | -.05 | 1 | | | | | | | | | | | |
| 7. Same Different RT | -.18 | -.05 | -.11 | -.15 | -.19 | .09 | 1 | | | | | | | | | | |
| 8. Synonym Accuracy | .26* | -.13 | .12 | .07 | .17 | .16 | .02 | 1 | | | | | | | | | |
| 9. Synonym RT | -.24^ | -.29* | -.11 | -.22^ | -.38** | -.03 | .51*** | .18 | 1 | | | | | | | | |
| 10. Hypernym Accuracy | .20 | .01 | .18 | .00 | .05 | -.17 | .14 | -.08 | .02 | 1 | | | | | | | |
| 11. Hypernym RT | -.28* | -.12 | -.18 | -.21 | -.30** | .11 | .48*** | .09 | .73*** | -.14 | 1 | | | | | | |
| 12. Free Association | .30* | .05 | .33** | .33** | .30* | -.25^ | -.12 | .21 | -.21 | -.14 | -.19 | 1 | | | | | |
| 13. Odd one out Accuracy | .34** | .26* | .19 | .34** | .05 | -.07 | -.08 | .02 | -.26* | .21 | -.39** | .27* | 1 | | | | |
| 14. Odd one out RT | -.09 | -.14 | -.04 | -.03 | -.26* | .07 | .15 | .06 | .32** | .09 | .52*** | -.17 | -.40*** | 1 | | | |
| 15. CoCo Errors | -.35** | -.23^ | -.06 | -.37** | -.10 | -.10 | -.01 | -.21 | .28* | -.18 | .08 | -.07 | -.32** | -.04 | 1 | | |
| 16. Mean RTs | -.32** | -.40** | -.27* | -.26* | -.28* | .10 | .34** | .26* | .68*** | -.05 | .59*** | -.27* | -.48*** | .33** | .26* | 1 | |
| 17. Interaction Term | .11 | -.01 | .16 | .11 | .21 | .19 | -.14 | .21 | -.22^ | -.10 | -.10 | .42*** | .12 | .03 | -.14 | -.21 | 1 |

Note: ^ $p < .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$.

The only variable that shared a significant correlation with the interaction effect was performance on the free association task ($r = .42, p = .001$). The only other variable that shared any correlation with the interaction effect was synonym RT, a relationship that approached statistical significance ($r = -.22, p < .08$). This measure comprised mean reaction times for correct synonym judgements, and reflected the speed with which participants were able to access the meaning of two words and to evaluate the similarity of those meanings.

In order to assess the contributions from each of these variables to the size of the interaction between coherence and cohesion, a stepwise hierarchical multiple regression with absolute interaction values as the dependent variable was conducted. It should be noted that the correlation between RTs in the synonym judgement task and interaction effect sizes may be an artefact of efficient judgement processes that are independent of semantic access. To account for this possibility, RTs from the discrimination task that did not necessitate access to semantics (the same/different task) were entered into the model in the first step of the analysis. RTs from the synonym judgement task were entered in the 2nd step, followed by the number of items produced in the free association task. This analysis is presented in Table 17.

Table 17. Stepwise hierarchical regression with absolute interaction value as the dependent variable (n=58)

| | <i>B</i> | <i>SE B</i> | β | ΔR^2 |
|----------------------|----------|-------------|---------------------|--------------------|
| Step 1 | | | | |
| Constant | .513 | .284 | | |
| Same/Different RTs | .000 | .000 | -.162 ^{ns} | .026 ^{ns} |
| Step 2 | | | | |
| Constant | .585 | .297 | | |
| Same/Different RTs | .000 | .001 | -.091 ^{ns} | |
| Synonym Judgment RTs | .000 | .000 | -.131 ^{ns} | .012 ^{ns} |
| Step 3 | | | | |
| Constant | -.126 | .356 | | |
| Same/Different RTs | .000 | .000 | -.084 ^{ns} | |
| Synonym Judgment RTs | -6.251 | .000 | -.052 ^{ns} | |
| Free Association | .007 | .002 | .396* | .15* |

* $p < .01$; [^] $p < .10$.

As Table 17 depicts, the only significant predictor of variance in the interaction values is performance on the free association task, which accounts for 15% of the variance in step 3. It was also the only significant predictor in the final model once the variance accounted for by the other variables was held constant. When the variance in interaction effect sizes that could be due to judgement speed irrespective to semantic access was partialled out in step 1, RTs from the synonym judgement task failed to predict any significant proportion of the variance in interaction effect sizes, accounting for just 1% of

the variance. So, it is likely that the moderate relationship between synonym RTs was dependent upon judgement speed rather than speed of semantic access per se.

5.4.4 Conclusions

The interaction between coherence and cohesion in was replicated in both the error and the standardised reading time data, demonstrating that this effect is a robust one. Contrary to the findings of Experiment 3, there was no 3-way interaction between coherence, cohesion and working memory group in the standardised reading time data. In addition, working memory capacity did not predict variance in the size of the interaction effect. This disparity is addressed in the general discussion section (see below).

The only variable that predicted variance in the size of the interaction between coherence and cohesion in the standardised reading time data was the total number of items produced in the free association task. The task required participants to produce words in response to a stimulus word, and is best described as a test of verbal fluency. Verbal fluency is often included in neuropsychological test batteries as an assessment of executive functioning (e.g. Ferstl et al., 2002). Executive function is implicated in goal directed behaviour, including holding and manipulating information in working memory and, critically, is believed to be involved in comprehension monitoring (e.g. Sesma et al., 2009). None of the other semantic tasks predicted variance in the size of the interaction effect, suggesting that this effect is not driven by access to semantics and cannot accurately be described as an 'elaboration' effect. Rather, the data suggest that this monitoring effect is driven by executive functioning.

However, performance on the free association task relies heavily on semantic ability and vocabulary: If subjects could not form a meaningful representation of the stimulus item and could not produce words that are semantically related to that representation then they performed poorly on the task. The task also shares significant correlations with vocabulary, the National Adult Reading Test and decoding skill (all r s $>.30$. See Table 16), all of which tap lexical ability. Thus, while the interaction effect may reflect comprehension monitoring, this monitoring appears to be related to the ability to form semantic representations. Comprehension monitoring is triggered by the presence of an atypical sentence pair, and this trigger engages the reader in a process of semantic elaboration; linking the properties of the text with world knowledge. The fact that working memory is not involved in this process suggests that comprehension monitoring is seated in language abilities rather than working memory capacity per se.

The present results do not accord well with the capacity constrained model of comprehension skill (Just and Carpenter, 1992). This account would have predicted a direct relationship between the interaction effect and working memory, over and above the language and semantic tasks. The lexical quality hypothesis would have predicted a relationship between working memory and the interaction effect, but that this would be mediated by decoding ability. This pattern of results is not demonstrated by the present data.

The present results are best explained by the long-term working memory model of Ericsson and Kintsch (1995), which correctly predicted that the interaction effect would be predicted by access to semantic knowledge. The fact that the free association task

correlates with measures of linguistic ability and verbal intelligence provides very weak evidence in favour of the distributed learning account of MacDonald and Christiansen (2002). More substantive support for this perspective would have been obtained had these tasks shared direct relationships with the interaction effect, however.

5.4.5 Combined Analyses

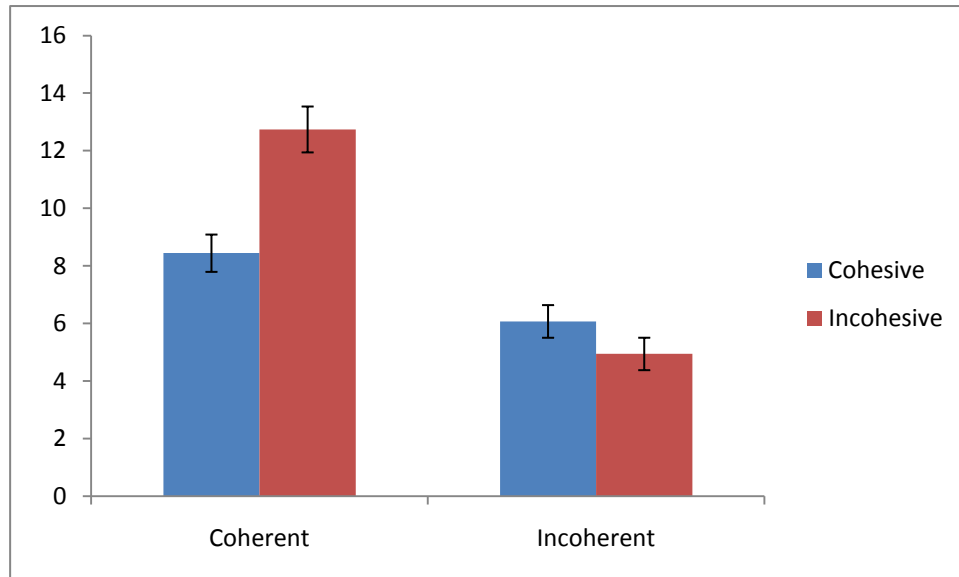
It is possible that the findings of Experiments 3 and 4 are not as discrepant as they appear. Rather, the relationship between working memory and the interaction effect in Experiment 4 may be weaker than that shown in Experiment 3. To test for this, the correlation coefficients between working memory and the interaction effect in studies 3 and 4 were compared using Fisher's Z transformation (Fisher, 1915). This analysis revealed that the two correlation coefficients did not differ to a statistically significant degree ($z = 1.18, p = .24$). As a consequence, the datasets were combined and analysed as a whole. These analyses are presented below.

Table 18. Descriptive statistics for the individual difference variables across subjects in Experiments 3 and 4 ($n = 115$)

| | Mean | SD |
|------------------------------|--------|-------|
| ND Comprehension | 235.52 | 14.33 |
| ND Vocabulary | 242.60 | 11.74 |
| Working Memory | 4.47 | 0.98 |
| Decoding composite(z-scores) | 0.00 | 0.84 |

Table 19. Accuracy and reading times for all participants in all conditions (mean(sd); n=115)

| | Coherent | | Incoherent | |
|--|--------------|-------------|-------------|--------------|
| | Cohesive | Incohesive | Cohesive | Incohesive |
| Percent error: | 8.44(6.95) | 12.74(8.55) | 6.07 (6.08) | 4.94(6.04) |
| Target sentence RTs (ms): | 1981(510) | 1858(488) | 2056(510) | 1805(466) |
| Std. Residuals of target sentence RTs: | -.0372(.220) | .0357(.277) | .1026(.228) | -.0329(.325) |

Figure 10. Mean percentage error rates for all conditions (n = 115). Error bars represent one standard error above and below the mean

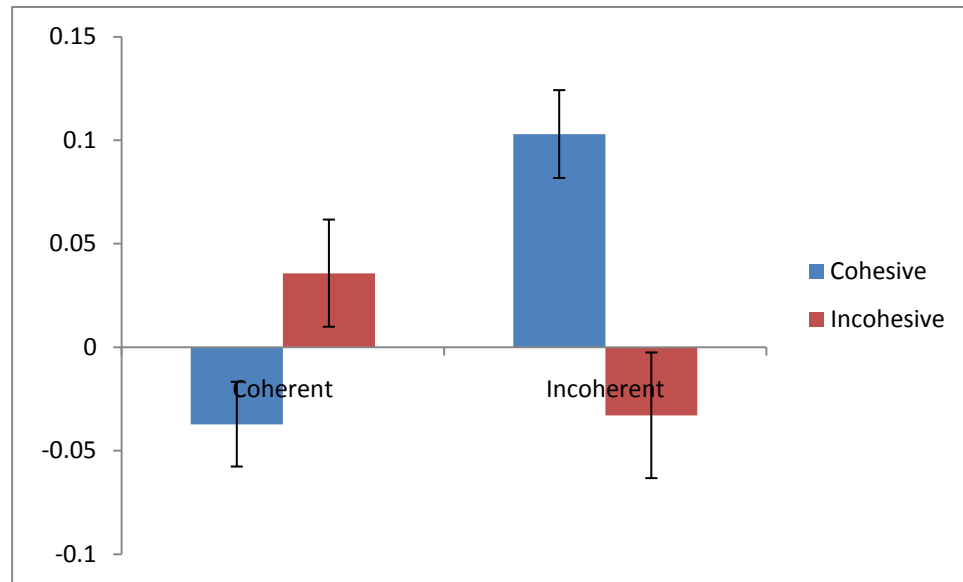
The accuracy data were subjected to a 2 x 2 analysis of variance. There was a main effect of coherence ($F(1, 114) = 46.13, p < .001$). Participants were more accurate on incoherent trials than they were on coherent trials (see Figure 10). There was a main effect of cohesion ($F(1, 114) = 10.27, p = .001$), showing that participants' judgements of coherence were, overall, more accurate for sentence pairs that were bound with cohesive ties. There was also an interaction between coherence and cohesion ($F(1, 114) = 26.50, p < .001$). This interaction was mainly driven by

differences in the coherent condition ($t(114) = 5.19, p < .001$), although the differences in the incoherent condition approached statistical significance ($t(113) = 1.89, p = .062$).

For the analysis of RTs, incorrect trials were excluded (872/10860; 8.3%). Mean reading times for each participant were then calculated for each participant across conditions. Trials that were 2.5 standard deviations above their mean RT were replaced with a cut off value (i.e. 2.5SD above their mean reading time). In total, 9975 trials were used for the analysis of reading times, 277 of which had been replaced with each individual's 2.5SD cutoff (2.8% in total). Just under ninety two percent of trials were retained for analysis.

As cohesive target sentences were longer than incohesive target sentences, it was necessary to factor-out sentence length for these analyses. For each participant separately, the reading times for target sentences were regressed on the number of characters (including spaces) in each sentence (i.e. sentence length). The standardised residuals that resulted from this procedure were then used for the following analyses. A graphical representation of the mean reading time residuals in the 4 conditions is presented in Figure 11.

Figure 11. Mean standardised residuals of target sentence reading times. Error bars represent one standard error above or below the mean ($n = 115$).



The mean standardised residuals were then used in a 2 (coherence) x 2 (cohesion) ANOVA. The interaction between coherence and cohesion depicted in Figure 11 was significant ($F(1, 114) = 28.94, p < .001$). Again, cohesive ties facilitated participant's comprehension of coherent sentence pairs, whereas their presence in incoherent sentence pairs impeded their comprehension. This impedance may stem from the fact that the coherence gap was contradicted by the presence of a coherence marker, causing readers to consider these cases in more depth and read them for relatively longer.

The intention of this combined analysis was to ascertain whether the interaction between coherence, cohesion and working memory group shown in Experiment 3 would hold once those data were combined with data from Experiment 4. Based on performance on the sentence span task, the participants were placed into 'low' and 'high' working memory groups. Those who

recalled 4 or fewer items in the sentence span task were placed into the 'low working memory' group, and those who recalled either 5 or 6 items were placed in the 'high working memory' group. The component reading characteristics of these two groups are presented in Table 20. Significant differences between the groups (as assessed by independent samples t-tests) are highlighted with asterisks.

Table 20. Component reading skills of the two working memory groups (mean(sd))

| | Working Memory Group | |
|-------------------------------------|----------------------|------------------|
| | High (n = 58) | Low (n = 57) |
| Mean ND comprehension scaled score: | 237.72 (13.13)^ | 233.28 (15.24)^ |
| Mean ND vocabulary scaled score: | 245.47 (9.79)** | 239.68 (12.87)** |
| Mean working memory (max=6): | 5.31 (0.47)*** | 3.61 (0.49)*** |
| Mean decoding composite (Z-scores): | .1501 (0.58)* | -.1527 (1.01)* |

*** $p < .001$; ** $p < .01$; * $p = .05$; ^ $p < .10$

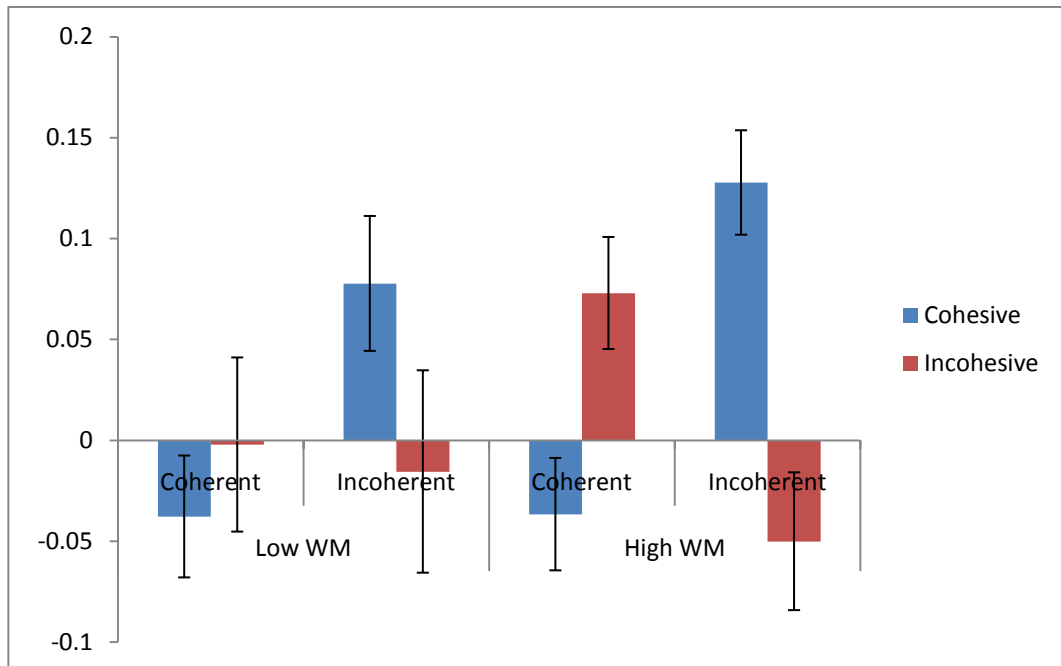
As shown in Table 20, the high working memory group performed marginally better than the low working memory group on the comprehension component of the Nelson Denny reading test ($t(113) = 1.68, p = .097$). They performed significantly better than the low working memory group on the vocabulary component of the Nelson Denny reading test ($t(113) = 2.71, p = .008$); the sentence span task ($t(113) = 18.99, p < .001$); and on the decoding composite ($t(113) = 1.97, p = .052$).

A 2 (coherence) x 2 (cohesion) x 2 (working memory group) mixed-design ANOVA, with working memory group as the between subjects factor, was conducted on the length-corrected residuals of the reading time data. There was no interaction between coherence and working memory group by subjects ($F_1(1, 113) = .235, p = .63$) or by items ($F_2(1, 2274) = 0.16, p = .69$), and there was no interaction between cohesion and working memory group by subjects ($F_1(1, 113) =$

.019, $p=.89$) or by items ($F_2(1, 2274) = 0.72, p = .40$). The analysis did reveal, however, a significant 3-way interaction between coherence, cohesion and working memory group by subjects ($F_1(1, 113) = 4.292, p < .04$) and by items ($F_2(1, 2274) = 4.46, p < .04$).

A 2 (coherence) x 2 (cohesion) ANOVA was conducted on the standardised reading time residuals of the two groups separately. Neither group showed main effects of coherence or cohesion (all $F_s < 2$). The low working memory group showed a moderate interaction between coherence and cohesion by subjects ($F_1(1,56) = 5.09, p < .03$) and a strong interaction by items ($F_2(1, 1100) = 15.70, p < .001$). In contrast, the high working memory group showed a very strong interaction between coherence and cohesion by subjects ($F_1(1, 57) = 31.05, p < .001$) and by items ($F_2(1, 1174) = 52.01, p < .001$). See Figure 12 for a graphical representation of these findings.

Figure 12. Mean residuals of target sentence reading times for the high working memory group ($n=58$) and the low working memory group ($n = 57$). Error bars represent one standard error above and below the mean



As can be seen in Figure 12, both groups showed an interaction between coherence and cohesion, but group differences arose in the coherent condition. The residual RTs of the high WM group are significantly different between the coherent/cohesive and coherent/incohesive conditions ($t(57) = 2.70, p = .009$). This effect is not demonstrated by the low WM group, however ($t(56) = 1.057, p = .295$). Thus, when presented with sentences that cohere but in which coherence is not reinforced by the presence of a cohesive tie, both groups of readers show a relative slowing of target sentence RT, but this effect is much more pronounced for high WM participants than it is for low WM participants.

To assess the contribution of the component variables to the size of the interaction between coherence and cohesion, a step-wise hierarchical multiple regression analysis was conducted using the interaction value of the standardised residuals as the dependent variable. The interaction value was ascertained in the same manner as in Experiments 3 and 4.

Table 21. Pearson's correlation coefficients between component variables and the interaction values

| | ND Comprehension | ND Vocabulary | Working Memory | Decoding Composite | Interaction Value |
|--------------------|------------------|---------------|----------------|--------------------|-------------------|
| ND Comprehension | 1 | | | | |
| ND Vocabulary | .54** | 1 | | | |
| Working Memory | .11 | .22* | 1 | | |
| Decoding Composite | .42** | .57** | .21* | 1 | |
| Interaction Value | .06 | -.01 | .24** | .15 | 1 |

** $p < .01$; * $p < .05$.

The only variable that shared a significant correlation with the interaction value was working memory. This variable was used as the sole predictor of the interaction value the regression analysis presented in Table 22.

Table 22. Regression analysis with absolute interaction value as the dependent variable

| | <i>B</i> | <i>SE B</i> | β | ΔR^2 |
|----------------|----------|-------------|---------|--------------|
| Constant | -.251 | .178 | | |
| Working Memory | .103 | .039 | .241 | .058* |

* $p = .009$

As can be seen in Table 22, working memory accounts for just under 6% of the variance in interaction effect sizes.

5.5 General Discussion

The combined analysis presented in Section 5.4.5 demonstrates that the relationship between working memory and the coherence x cohesion interaction effect, though weaker in Experiment 4, was present in data obtained from both experiments. The most obvious explanation of the difference between Experiments 3 and 4 would be differences between the working memory groups. That is, it may be that the 'low' and 'high' working memory groups were more extreme in Experiment 3. As can be seen in Table 23, however, the two groups were relatively similar across the two studies:

Table 13. Span characteristics of the low and high working memory groups in Experiments 3 and 4

| Experiment No. | 3 items recalled (low) | 4 items recalled (low) | 5 items recalled (high) | 6 items recalled (high) |
|----------------|---------------------------|---------------------------|----------------------------|----------------------------|
| Three | N = 8 | N = 16 | N = 20 | N = 8 |
| Four | N = 14 | N = 19 | N = 20 | N = 10 |

It is plausible that the sentence span task may not be discriminating between subjects as best it could. There is evidence to suggest that the sentence span task may be quite unreliable. Waters and Caplan (1996) tested 44 college students on the reading span task over 2 sessions separated by 3 months. Span performance at time 1 and time 2 shared a correlation coefficient of just .41. At time 1, Waters and Caplan divided subjects into low, medium and high-span groups (low = 2 - 2.5 items recalled; medium = 3 - 3.5 items recalled; high = 4 or more items recalled). In session two, 44% of the subjects changed in terms of their classification group. Of the 3 high-span subjects who shifted category, 2 became medium-span and one became low-span. Of the 9 medium span subjects who shifted category, 5 became high-span and 4 became low-span. Of the 6 low-span subjects who shifted category, 3 became medium-span and 3 became high-span (Waters & Caplan, 1996).

In the present data, the sentence span task shows inconsistent relationships with measures other than the interaction effect. In Experiment 3, performance on the sentence span task was significantly correlated with reading comprehension skill ($r = .41$, $p < .001$. See Table 12) but in Experiment 4, this relationship was essentially nonexistent (r

= $-.13$, $p > .05$. See Table 16). Fisher's transformation test revealed that the relations between working memory and ND comprehension in the two experiments were significantly different from one another ($z = 2.88$, $p = .004$).

The span task and the comprehension component of the Nelson Denny Reading Test were administered in exactly the same way and at the same time-points in both experiments (the end of session 1 for the Nelson Denny; the beginning of session 2 for the sentence span task).

The span task shared a significant relationship with the decoding composite in Experiment 3 ($r = .31$, $p < .05$), but the same relationship did not hold in Experiment 4 ($r = .15$, $p > .05$). However, Fisher's transformation test showed that these correlation coefficients did not differ to a statistically significant degree ($z = 0.86$, $p = .39$). The relationship between the span task and the vocabulary component of the Nelson Denny reading test still held in Experiment 4, but was more modest than in Experiment 3 ($r = .31$, $p < .05$ in Experiment 3; $r = .26$, $p < .05$ in Experiment 4). Fisher's transformation showed that this difference was not significant ($z = 0.28$, $p = .78$).

The findings of Experiment 4 should be interpreted with some caution; working memory capacity is still related to the size of the interaction between coherence and cohesion when data are collapsed across experiments, and the assertion that the effect reflects a comprehension monitoring process is still plausible. However, the fact that performance on the verbal fluency task predicts a significant portion of the variance in the effect (Experiment 4) indicates that the process is at least partly dependent on semantic ability and access to world knowledge.

The standardised residuals of the reading time data varied as a function of coherence and cohesion (i.e. consistency), and this variance was predicted by differences in working memory capacity (Experiments 3 and 4) and semantic fluency (Experiment 4). In both experiments, readers with high working memory capacities showed larger interactions between coherence and cohesion. These differences were most pronounced in the coherent/incohesive and incoherent/cohesive conditions. In these trials, high capacity readers showed large effect sizes in a positive direction, equating to slower reading times of the target sentences. High-capacity readers may be more sensitive to the function of cohesive ties; when ties linked incoherent sentences (incoherent/cohesive), this may have triggered an inconsistency warning, resulting in a longer and more pronounced appraisal of the target sentence. When a cohesive tie was *missing* between coherent sentences (coherent/incohesive), this signalled an inconsistency, again resulting in a greater degree of appraisal of the target sentence.

In Experiment 4, semantic fluency (the free association task) accounted for unique variance in the size of the interaction between coherence and cohesion. This finding can be interpreted with reference to the semantic deficit hypothesis (e.g. Nation & Snowling, 1999; Landi & Perfetti, 2007), which states that reading comprehension difficulties result from deficits in semantic processing. Nation and Snowling (1999) found that poor comprehenders were less sensitive to abstract semantic relations than were normal readers. Perfetti and Landi (2007) demonstrated that skilled and less-skilled comprehenders showed different ERPs in response to semantically related word pairs. As stated in the introduction, less-skilled comprehenders showed weaker meaning-

congruence effects in Landi & Perfetti's (2007) study, suggesting a deficit in semantic processing. The findings of Experiment 4 indicate that semantic processing abilities determine the degree to which readers will reflect upon pragmatic relations between sentences: Weaker semantic skills predict less reflection, and stronger semantic skills predict more reflection.

The findings of Experiments 3 and 4 also shed some light on models of working memory. The 3-way interaction between coherence, cohesion and working memory group in the standardised residuals of the reading time data can be interpreted with reference to the capacity-constrained model of comprehension (Just & Carpenter, 1992): High-capacity readers were able to devote more cognitive resources to their evaluations of coherence than were low-capacity readers. Experiment 4 may go some way to elucidating the mechanism that underlies this evaluative process: The degree to which readers will reflect upon pragmatic relations between sentences varies as a function of semantic fluency. This finding fits with both the long-term working memory model of Ericsson and Kintsch (1995), and the connectionist account of working memory proposed by MacDonald and Christiansen (2002). Both of these models emphasise the role of experience and skill in relation to reading processes. Ericsson and Kintsch (1995) argue that individuals draw on acquired knowledge through systems of 'retrieval structures' that encode information into long term memory in a retrievable form. These structures, activated by incoming information, act to retrieve pertinent information from long term memory into a temporary storage space where it can be used to execute a task (Ericsson & Kintsch, 1995). Similarly, MacDonald and Christiansen (2002) propose that concepts

that are frequently activated (through language experience) have a higher level of resting activation, meaning that minimal activation will result in an instantiation of those concepts into the current task. The number of items participants were able to produce in the semantic fluency task may be indicative of the quality of these retrieval structures and/or the level resting activation of concept nodes. That is well-specified, high-quality retrieval structures or nodes will be activated on the basis of stimulus items, with more information from long term memory being retrieved in response to them (Ericsson & Kintsch, 1999), or more semantic associates being instantiated by via spreading activation across a semantic network (MacDonald & Christensen, 2002). With reference to the CI model (Kintsch, 1988; 1998), the context sentence may activate relevant retrieval structures/concept nodes, drawing information from long term memory which is then compared to the activations instantiated by the target sentence. Those who have well-specified retrieval structures or concept nodes will activate more information from the context and target sentences. When lexical information indicates the presence of a pragmatic relation where none exists (the incoherent/cohesive condition), or the lack of lexical information does not reinforce a pragmatic relationship (the coherent/incohesive condition), those who have activated more information from long term memory will reflect upon the relations between the context and target sentences to a greater degree than those who have activated less information. Importantly, the conditions that produced large effect sizes in the coherence and cohesion paradigm are quite atypical: It is rare to encounter coherent sentences that are not linked with a cohesive tie (coherent/incohesive), and it is very rare to encounter incoherent sentences that are

erroneously linked with a cohesive tie (the incoherent/cohesive condition). Thus, readers who have a greater degree of linguistic experience not only show superior semantic processing skills and a greater degree of concept activation/memory retrieval, they also show a greater sensitivity to lexical cohesion. When they are confronted with atypical sentence constructions in which cohesive ties misleadingly link incoherent sentences or cohesive ties are conspicuously absent from coherent sentences, they reflect upon the pragmatic relations between the two sentences (given by concept activation/long term memory retrieval), resulting in relative increases in processing time.

These conclusions need to be taken tentatively, however. The capacity constrained model of comprehension skill (Just & carpenter, 1992), would have predicted a much stronger relationship between working memory and the interaction effect, one that would have been robust in both experiments, which was not the case. Further, if the distributed learning account of MacDonald and Christiansen (2002) is entirely correct, then the purer measures of language proficiency and intelligence (decoding skill and the NART) would have been more directly related to the interaction effect. Last, if access to world knowledge in long term memory drives the relation between working memory and comprehension skill, then the other measures of semantic ability should have also been related to variance in the interaction effect.

The findings of these studies also lend no support to the lexical quality hypothesis, which would have predicted a relationship between decoding ability and the interaction effect that is mediated by working memory capacity. No such relationship was revealed, however.

6 Summary and General Discussion

6.1 Chapter Overview

The aim of this thesis was to investigate some of the processes that contribute to the effective comprehension of text in an adult population. The relation between gist-based memory processes and reading comprehension skill was explored in Experiment 1, component skills that contributed to performance in an off-line test of reading comprehension skill were investigated in Experiment 2, and skills that contributed to integration processes in a coherence judgement task were explored in Experiments 3 and 4.

A summary of the findings of the experiments reported in this thesis is first presented. Subsequent sections discuss the empirical findings in relation to gist-based memory processes, component skills of comprehension, coherence building and the relationship between working memory, lexical quality, semantic fluency and reading comprehension skill. It is argued that a comprehensive model of text comprehension in adults needs to consider the relative influences of these skills on both on-line and off-line comprehension processes.

6.2 Summary of Findings

Experiment 1 explored the relation between gist-based memory processes and reading comprehension skill using the DRM paradigm. The hypothesis was that adults with reading comprehension weaknesses would show reduced rates of false recall and

recognition, indicative of deficits in semantic processing and problems extracting the gist meaning from incoming stimuli. Weekes et al. (2008) found that children with comprehension difficulties produced fewer extra-list intrusions than did good comprehenders, suggesting that poor comprehenders had difficulty extracting the central theme or gist from the semantically related word lists; a deficit that was attributed to weakness in semantic processing and memory. Experiment 1 demonstrated that this effect was not replicable in an adult population.

In Experiment 2, measures of vocabulary, word-level skill, working memory and verbal intelligence were used as predictors of comprehension skill in stepwise, multiple regression analyses. Word-level skill and vocabulary were found to be significant predictors of variance in reading comprehension skill. Verbal intelligence also accounted for a small yet near-significant proportion of the variance in comprehension skill. Although there was a significant correlation between working memory capacity and performance on the comprehension component of the Nelson Denny Reading Test, this relationship failed to hold in the multiple regression analyses once the contributions of vocabulary and word-level ability had been accounted for.

Experiments 3 and 4 explored the processes involved in on-line reading using a comprehension task that demanded integration. Subjects took part in a coherence judgement task (Ferstl, Guthke & von Cramon, 2002; Ferstl, 2006) in which they had to verify whether two sentences cohered with one another or not. Four conditions that resulted from crossing coherence and cohesion (i.e. the presence of a lexical connection), were used: Coherent and cohesive (where sentences cohered, and a cohesive tie made

their coherence explicit); coherent and incohesive (where sentences cohered, but coherence had to be inferred on the basis of pragmatic information rather than being explicitly signalled); incoherent and cohesive (where sentences that do not cohere were erroneously linked with a cohesive tie); and incoherent and incohesive (where sentences did not cohere, and were not erroneously linked with a cohesive tie). Experiment 3 demonstrated an interaction between coherence and cohesion, both in the error data and the length-corrected reading time data. When sentences cohered and their coherence was reinforced by a cohesive tie, subjects read them faster and made fewer errors on them. When sentences cohered but coherence was not reinforced with a cohesive tie, subjects read them more slowly and erroneously judged them to be incoherent more frequently. When sentences did not cohere but coherence was falsely indicated by the presence of a cohesive tie, subjects again read them more slowly and frequently misjudged them as coherent. Last, sentences that did not cohere and that were not falsely linked with cohesive ties were read more quickly and were less likely to be misjudged as coherent.

In the reading-time data, high working memory subjects demonstrated a greater sensitivity to the experimental manipulations, showing larger interactions between coherence and cohesion than the low working memory subjects. Interestingly, the same pattern did not manifest in the error data – working memory was not related to coherence judgements. This result was interpreted as a monitoring effect, elicited by atypical sentences that did not adhere to usual constructions.

Experiment Four sought to explore the semantic deficit hypothesis in relation to this effect, with the proposal that efficient semantic processes rather than working memory capacity contributed to variance in the size of the interaction between coherence and cohesion. Performance on a semantic fluency task was found to predict unique variance in the size of the interaction effect, over and above that accounted for by working memory capacity and any other individual difference variable. This finding suggests that the effect may be explained by language processing skills based in semantics than by working memory capacity per se.

6.3 Theoretical Implications

6.3.1 Gist-Based Memory Processes

It was argued that adults with comprehension weaknesses do not have difficulties activating semantic associates in the DRM paradigm and, as discussed in Chapter 2, this effect may be a proxy for their ability to activate items during the process of construction. Instead, comprehension problems in adulthood may be characterised by problems involving integration processes during text comprehension (Long & Chong, 2002).

There is plenty of evidence to suggest that adults with reading comprehension difficulties have deficits in semantic access and word-level processes (e.g. Perfetti et al., 2008). The findings of Weekes et al. (2008) suggested that, in children, text comprehension difficulties are related to problems extracting the gist meaning of incoming stimuli, which may be symptomatic of difficulties extracting the gist or global theme from texts. The Construction-Integration model proposes that such gist-extraction

is critical for the formation of a macrostructure of a text by allowing the propositions given by texts to be organised hierarchically (Kintsch, 1998), and it has been shown that children with reading comprehension impairments show difficulties in stating the main ideas and themes of given texts (Yuill & Oakhill, 1991).

It is known that children become more adept at global gist extraction as they get older (Odegard, Holliday, Brainerd, & Reyna, 2008), and there is evidence to suggest that when children are very knowledgeable about a topic, they are better able to extract gist meanings from lists of items related to one another on the basis of that topic. For instance, some evidence indicates that children who are dinosaur experts begin to solve class inclusion problems dealing with dinosaurs at an earlier age than do children who lack this expertise (Ni, 1998). The literature review discussed the importance of background knowledge in reading comprehension (e.g. Bransford & Franks, 1972; Long et al., 2008), and it may be that the poor comprehenders in the Weekes et al. (2008) study simply had poorer knowledge about the domains to which the word lists applied. Since items in the DRM paradigm refer to everyday, commonly experienced activities, adults' knowledge of these domains should be fairly well developed, meaning that extracting the gist from items that are related to these domains is relatively simple, and does not distinguish good and poor comprehenders. This explanation does not, however, fit with the findings of Cain and colleagues (Cain et al., 2001), who found that good and poor comprehenders with comparable levels of explicit knowledge about a fictional world still differed in their ability to use that information to make inferences. Another possibility, alluded to in the discussion section of Experiment 1, is that adults with weak comprehension skill do not

have difficulties activating semantic associates in the process of construction. Rather, their deficits may be related to integrating those associates into a meaningful text-base level of representation. Though there is evidence that this may be correct (see Long & Chong, 2002), the present data do not warrant a full subscription to this interpretation.

A simpler explanation is that the adult readers classified as ‘poor comprehenders’ in Experiment 1 were not as extreme a group as the poor comprehenders in the Weekes et al. (2008) study. The Weekes et al. study classified poor comprehenders as those whose estimated comprehension age was more than 1 year below their estimated reading accuracy age and 1 year below their chronological age (as in Oakhill, 1982). In contrast, the poor comprehenders in Experiment 1 were those whose comprehension score fell below the sample median. Furthermore, the subjects used in Experiment 1 were all undergraduate students at the University of Sussex and were in possession of reading skills that permitted them entry to higher education. In retrospect, a broader sample of readers from a variety of educational backgrounds should have been used for a fuller investigation of gist-based processing in adult readers.

6.3.2 Component Skills

Experiment 2 found that vocabulary and word-level skills each accounted for unique portions of variance in reading comprehension skill and that verbal intelligence accounted for a small yet non-significant amount of variance. Although it is difficult and potentially misleading to compare across different types of statistical analysis, is it interesting to note that in Experiment 2, working memory shared a significant correlation

with comprehension skill (see Tables 4 & 5), but when the contributions of vocabulary and decoding skill were accounted for in the multiple regression analysis, working memory failed to predict any unique variance in comprehension skill (see Tables 6 & 7). This pattern of results fits, in part, with the Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007).

The lexical quality hypothesis holds that lexical access needs to be accurate (i.e. knowing word meanings), fluent (i.e. efficient), and rich (lexical items are connected to associates through an elaborate knowledge net). Laboured or 'sluggish' word identification constrains the reading process by: (1) impeding syntactic analysis (e.g. tense and plurality cannot be properly encoded if semantic specificity cannot be established); (2) restricting the number of semantic associates that are activated and/or integrated with the current mental representation during the process of construction and integration; and (3) Monopolising working memory resources that would otherwise be assigned to higher-order comprehension processes such as sentence integration and inference generation. The first two of these proposals predict a *direct* relationship between word-level ability and reading comprehension skill, an effect that would presumably operate during the construction of the surface and text-base levels of text representation. The last of these predicts an *indirect* relationship that is mediated by working memory capacity or function that would operate at the discourse-level of text representation. The pattern of results presented in Chapter 3 support the first two of these proposals, but fails to provide evidence for the third.

6.3.3 Working Memory as a Component of Reading Comprehension

Working memory shared a relationship with reading comprehension skill only when the influences of word-level skill and vocabulary were not accounted for. This finding accords, at least in part, with the *experience* based models of working memory; the Long-Term Working Memory Model of Ericsson and Kintsch (1995) and the connectionist account of working memory given by MacDonald and Christiansen (2002). MacDonald and Christiansen contend that processing capacity emerges from network architecture and experience, and that working memory is not a separate entity but is the network *itself*. Thus, in tasks such as the sentence span task (Daneman & Carpenter, 1980), performance reflects the language system's ability to efficiently process incoming information. The pattern of results in Experiment 2 accords with the idea that the relationship between working memory and comprehension skill reflects linguistic aptitude rather than general processing ability. As stated in the materials section of Experiment 1, the National Adult Reading Test provides a surrogate measure of verbal intelligence. The finding that verbal intelligence accounts for a small but near-significant proportion of the variance in comprehension skill adds some weight to the proposal that reading comprehension depends upon linguistic aptitude. However, the question as to whether working memory capacity reflects linguistic ability rather than general processing ability cannot be answered with the present data.

6.3.4 Coherence, Cohesion, Working Memory and Semantic Fluency

Experiment 3 showed that high working memory subjects were more sensitive to the manipulations of the coherence and cohesion paradigm than were low working memory subjects. This finding was interpreted as a monitoring effect, with the presence or absence of cohesive ties in the *atypical* conditions acting as a trigger for the metacognitive process of comprehension monitoring. Those participants with higher working memory capacities showed a greater sensitivity to these triggers by virtue of their having resources available to attend to them, an interpretation that accords with the capacity constrained model of comprehension of Just and Carpenter (1992). However, Experiment 4 found that the influence of working memory on the interaction between coherence and cohesion was markedly weaker than in Experiment 3, and it was only when data from the two experiments were combined that the effect was revealed. When the data of Experiment 4 were considered in isolation, it was performance on a semantic fluency task (the free association task) that accounted for a significant portion of the variance in the magnitude of the interaction.

Verbal fluency tasks have been used to assess executive function in a number of neuropsychological investigations (e.g. Monsch, Bondi, Butters, Salmon, Katzman, & Thal, 1992), but performance on *semantic* fluency tasks is thought to depend more on lexical knowledge and semantic memory organisation than on executive functioning (Ardila, Ostrosky-Solis, & Bernal, 2006). Semantic fluency tasks have been shown in Positron Emission Tomography studies to activate cortical regions that are more strongly associated with language processing (i.e. the temporal lobe) than regions that are

associated with executive functioning (e.g. the frontal lobe) (Warburton, Wise, Price, Weiller, Harad, Ramsay, & Frackowiak, 1996).

Both the long-term working memory theory of Ericsson and Kintsch (1995) and the connectionist account of working memory by MacDonald and Christiansen (2002) can go some way to explaining the disparate relation between task performance and working memory that was exposed by Experiments 3 and 4. If working memory capacity is driven by linguistic experience, as is proposed by both of these theories, then sentence span performance is really a proxy for linguistic ability. For high working memory readers, the atypical conditions will trigger a relative slowdown in processing speed because of their irregular structure: The language processing system has not had experience with sentence constructions of this type and as a consequence, will subject them to a deeper level of processing. However, no measure of linguistic experience was taken as part of this thesis. Furthermore, no measure of linguistic ability could account for any variance in the size of the interaction between coherence and cohesion. Although appealing, the proposal that the larger effect sizes shown by high working memory subjects can be attributed to language processing ability is not warranted by the present data.

As stated earlier in this thesis, McNamara and her colleagues (McNamara, Kintsch, Songer, & Kintsch, 1996; McNamara, 2001) have shown that readers with a high level of domain knowledge benefit most from texts that are less cohesive, presumably because they are forced to reflect upon background knowledge to integrate propositions in the text into a coherent mental representation. With reference to the CI model (Kintsch, 1988; 1998), performance on the fluency task may be related to the processes of construction

and integration that serve to develop a text-base level of text representation. Items presented to subjects in the fluency task presumably serve to activate associates in a manner that is akin to priming. These associates are then selected for production on the basis of their relation to the stimulus item, with highly associated items reaching a level of activation that triggers their vocalisation. It would be tempting to infer that performance on the fluency task may be indicative of greater general world knowledge or access to world knowledge. However, the stimulus items used in the fluency task were the critical distracters of the DRM paradigm, all of which are highly related to knowledge domains with which most, if not all, people in the western world have experience. Further, Experiment 1 suggested that both good and poor readers activate world knowledge to the same degree. So, it may be that all readers, regardless of comprehension skill, activate related items in the free association task, but those who have good integration abilities vocalise those related words.

6.4 Conclusions

Taken together, the findings presented in this thesis emphasise the importance of lexical-semantic processes in adult reading comprehension skill. This conclusion is warranted by the finding that vocabulary and word-level ability were predictive of comprehension skill in Experiment 2, and that performance on a semantic fluency task was predictive of integration processes in Experiment 4. Working memory capacity shared a tenuous relationship with reading comprehension skill in Experiment 2, and was inconsistent in its association with integration processes across Experiments 3 and 4. This

pattern of results was explained with reference to experience-based models of working memory, which posit that it is the efficiency with which the language system can deal with linguistic information that determines a reader's level of reading ability and comprehension skill. In Experiment 2, working memory shared a correlational relationship with comprehension ability that failed to hold when the contributions of vocabulary and word-level ability were accounted for. Similarly, in Experiment 3 working memory was predictive of the magnitude of the interaction between coherence and cohesion, but in Experiment 4 this variance was exclusively predicted by performance in a semantic fluency task. In both cases, linguistic abilities superseded working memory capacity. However, measures of linguistic experience or background knowledge were not taken as part of any experiment presented in this thesis, so the experience-based models of working memory and reading comprehension skill cannot be fully endorsed.

With reference to the CI model, the results of this thesis suggest that variation in adult comprehension skill can be attributed, at least partly, to differences in integration processes. Experiment 1 showed that adults with comprehension weaknesses do not have difficulties activating semantic associates in the DRM paradigm. This finding was explained with reference to background knowledge (i.e. all adults have the requisite domain knowledge to extract the gist meaning or theme from the semantically related word lists) and suggests that reading comprehension in adults is not related to problems in construction. Experiment 2 found that decoding ability and vocabulary were significant predictors of variance in comprehension skill. This suggested that lexical quality is related to comprehension processes in adult readers, but the fact that working memory failed to

predict any unique variance in comprehension skill and did not interact with the lexical tasks suggests that word-level skill does not affect comprehension processes by monopolising working memory resources.

The mixed pattern of results presented in Chapter 5 suggested that integration processes are dependent on efficient and fluent semantic access and . Experiment 3 showed that the magnitude of the interaction between coherence and cohesion was dependent on working memory capacity. This relationship did not hold as strongly in Experiment 4, where the interaction was predicted by performance on a semantic fluency task. This pattern of results was tentatively interpreted with reference to the findings of McNamara and colleagues (McNamara & Kintsch, 1996; McNamara et al., 1996; McNamara, 2001), and experience-based models of working memory and text comprehension (MacDonald & Christiansen, 2002; Ericsson & Kintsch, 1995).

McNamara and colleagues have demonstrated that breaks in coherence induce readers to reflect on prior knowledge to make inferences and fill-in coherence gaps. The results of Experiment 4 suggest that semantic fluency determines the degree to which subjects will reflect on prior knowledge in response to coherence gaps.

The connectionist and long-term working memory models would imply that readers' sensitivity to the manipulations of the coherence and cohesion paradigm is determined by their linguistic experience. Though appealing, no measures of language experience were taken as part of these experiments, and none of the linguistic aptitude measures accounted for any of the variance in the interaction between coherence and cohesion. As such, this suggestion cannot be substantiated by the present data.

6.5 Future Directions

Ambiguity remains with regard to the role of working memory and/or executive functions in integration, as working memory was found to be related to integration processes in Experiment 3 whereas semantic fluency was found to be related to integration processes in Experiment 4. Since both of these tasks are thought to rely on executive skills, it is difficult to ascertain the relative contributions of semantic ability, executive skills and experience-based linguistic aptitude to variance in the size of the interaction between coherence and cohesion. One way to dissociate these influences would be to re-run the coherence and cohesion paradigm and to administer non-linguistic measures of executive function along with measures of print exposure to give an indication of participants' reading experience.

One measure of executive functioning that does not depend upon language ability to the same extent as the Reading Span Task is the Operation Span Task (Turner & Engle, 1989). In this task, subjects receive an *operation* in equation format and have to state whether it is 'true' or 'false' (e.g. $8*2 = 15$ – 'false'), after this they are then presented with a to-be-remembered word (e.g. 'tree'). Set sizes increase from 2-5 items, meaning that the maximum number of operations and to-be-remembered words is 5. This task has been strongly implicated in the central executive functions of monitoring and updating (Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000). If sensitivity to the manipulations of the coherence and cohesion task is due to monitoring processes (as

suggested in Experiment 3) then performance on this task should predict unique variance in the size of the interaction between coherence and cohesion.

Print exposure can be estimated using tests such as the Author Recognition Test (ART), the Magazine Recognition Test (MRT) and the Title Recognition Test (TRT) (Stanovich & West, 1989). In these measures, subjects must recognise authors, magazines or book titles from a number of real authors, magazines and book titles embedded among foils. Accuracy on these tests has been shown to predict the literacy levels of readers in a variety of studies (e.g. Stanovich & Cunningham, 1992; West & Stanovich, 1991; Stanovich & West, 1989). If readers' sensitivity to the manipulations of the coherence and cohesion paradigm is due to reading experience rather than updating and monitoring operations of the central executive, then performance on these tasks should predict variance in the size of the interaction between coherence and cohesion, over and above that predicted by performance on the operation span task.

It may be that print exposure contributes to an efficient language processing system that affects both semantic processing and executive functions via a process of *reciprocal causation*. That is, it is possible that differences in exposure to print affect the development of cognitive processes and declarative knowledge bases in an interactive manner (see Stanovich & Cunningham, 1992). If so, then *all three* of the proposed measures (semantic fluency, operation span, and print exposure) should account for unique portions of variance in the magnitude of the interaction between coherence and cohesion.

6.6 Final Considerations

The aim of this thesis was to investigate the processes involved in reading comprehension in adults. The results emphasise the importance of vocabulary and sub-lexical skills to the reading process, and suggest that the role of working memory in reading comprehension is one that is intrinsically related to language processing skill.

Future research should try to dissociate the relative influences of language processing ability, background knowledge and executive processes on comprehension skill. If a clear depiction of text comprehension in adults is to be achieved, then the contributions of each of these skills to the reading process need to be isolated, and interrelations and dependencies among them need to be quantified.

References

- 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.
- Ardila, A., Ostrosky-Solis, F., & Bernal, B. (2006). Cognitive testing toward the future: The example of Semantic Verbal Fluency (ANIMALS). *International Journal of Psychology*, 41(5), 324-332.
- Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), *Recent advances in the psychology of learning and motivation*, 8, 647-667. New York: Academic Press.
- Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, 36 (3), 189-208.
- Baker, L., & Stein, N. (1981). The development of prose comprehension skills. In C.M Santa & B.L. Hayes (Eds.), *Children's Prose Comprehension: Research and Practice* (pp. 7-43). International Reading Association, Newark, Delaware.
- Bartlett, F. C. (1932). *Remembering: A study of experimental and social psychology*. Cambridge, England: Cambridge University Press.
- Beck, I.L., Perfetti, C.A., & McKeown, M.G. (1982). Effects of Long-Term Vocabulary Instruction on Lexical Access and Reading Comprehension. *Journal of Educational Psychology*, 74 (4), 506-521.
- Bell, L. C. & Perfetti, C.A. (1994). Reading Skill - Some Adult Comparisons. *Journal of Educational Psychology* 86(2), 244-255.
- Boesma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*. 5:9/10, 341-345.
- Brainerd, C. J., & Reyna, V. F. (2005). *The Science of False Memory*. New York: Oxford University Press.
- Bransford, J. D. and M. K. Johnson (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior* 11(6), 717-726.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities* 40 (3), 226-243.

Brown, J.I., Fischco, V.V., & Hanna, G. (1993). Nelson-Denny Reading Test. Chicago: Riverside.

Cain, K. & Oakhill, J.V. (1999). "Inference making ability and its relation to comprehension failure in young children." *Reading and Writing* 11(5-6): 489-503.

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

Catts, H.W. Fey, M.E., Zhang, X. and Tomblin, J.B. (1999). Language Basis of Reading and Reading Disabilities: Evidence From a Longitudinal Investigation. *Scientific Studies of Reading*, 3(4), 331-361.

Daneman & Carpenter, 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19(4), 450-466.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58, 17-22.

Ericsson, K. A., Kintsch, W. (1995). Long-Term Working-Memory. *Psychological Review*, 102 (2), 211-245

Ferstl, E. C., & von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Cognitive Brain Research*, 11(3), 325-340.

Ferstl, E.C. Guthke, T. & von Cramon, D.Y. (2002). Text comprehension after brain injury: Left prefrontal lesions affect inference processes. *Neuropsychology*, 16, 292-308.

Ferstl, E. C. (2006). Text comprehension in middle aged adults: Is there anything wrong? *Aging Neuropsychology and Cognition*, 13(1), 62-85.

Fisher, R.A. (1915). Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika* (Biometrika Trust) 10 (4), 507-521.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing Inferences During Narrative Text Comprehension. *Psychological Review*, 101(3), 371-395.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K.M. (2004). Integration of Word Meaning and World Knowledge in Language Comprehension. *Science*, 304, 438-441.

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

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 98, 122–149.

Kintsch, W. (1988). The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model, *Psychological Review*, 95, 163-182.

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

Kintsch W & van Dijk T A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-94.

Kirby, J.R. (2007). Reading Comprehension: Its Nature and Development. (pp. 1-8). *Encyclopedia of Language and Literacy Research Network*. Retrieved 14/7/10 from <http://literacyencyclopedia.ca/pdfs/topic.php?topId=227>

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences* 4(12), 463-470.

Kutas, M., & Hillyard, S. A. (1980). Reading Senseless Sentences - Brain Potentials Reflect Semantic Incongruity. *Science*, 207(4427), 203-205.

Landi, N. & Perfetti, C. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language* 102(1): 30-45.

Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology-Learning Memory and Cognition*, 27(6), 1424-1429.

Long, D.L., Johns, C.L. & Morris, P.E. (2006). Comprehension Ability in Mature Readers. *Handbook of Psycholinguistics: 2nd Edition*. Elsevier Inc. ISBN: 0-12-369374-8.

Long, D. L., Oppy, B. J., & Seely, M. R.(1997). Individual differences in readers' sentence- and text-level representations. *Journal of Memory and Language*, 36(1), 129-145.

Long, D. L., Prat, C., Johns, C., Morris, P., & Jonathan, E. (2008). The importance of knowledge in vivid text memory: An individual-differences investigation of recollection and familiarity. *Psychonomic Bulletin & Review*, 15(3), 604-609.

Long & Prat, 2002. Memory for Star Trek: The role of prior knowledge in recognition revisited. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 1073-1082.

Long, D. L., Wilson, J., Hurley, R., Prat, C. S. (2006). Assessing text representations with recognition: The interaction of domain knowledge and text coherence. *Journal of Experimental Psychology-Learning Memory and Cognition*, 32(4), 816-827.

Lovett, M.C., Reder, L.M. & Lebiere, C. (1999). Modelling Working Memory in a Unified Architecture: An ACT-R Perspective. In Miyake, A. & Shah, P. (1999) *Models of Working Memory, Mechanisms of Active Maintenance and Executive Control*. Cambridge University Press.

MacDonald, M. C., Christiansen, M. H. (2002). Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, 109(1), 35-74.

McCallum, S. (2002). Phonemic Awareness in Literate Adults. PhD Dissertation. Department of Psychology, The University of Sussex, Falmer, Brighton.

McKoon, G., Ratcliff, R. (1992). Inference during Reading. *Psychological Review*, 99(3), 440-466.

McNamara, D.S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, 55, 51-62.

McNamara, D. S. & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes* 2(3): 247 - 288.

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

McNamara, D.S. & Magliano, J.P. (2009). Towards a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation*. New York, NY: Elsevier Science.

Miller, A. R., Baratta, C., Wynveen, C., & Rosenfeld, J. P. (2001). P300 latency, but not amplitude or topography, distinguishes between true and false recognition. *Journal of Experimental Psychology-Learning Memory and Cognition*, 27(2), 354-361.

- Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100.
- Monsch, A.U., Bondi, M.W., Butters, N., Salmon, D.P., Katzman, R., & Thal, L.J. (1992). Comparisons of Verbal Fluency Tasks in the Detection of Dementia of the Alzheimer Type. *Archives of Neurology*, 49(12), 1253-1258.
- Nation, K. & Snowling, M.J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: evidence from semantic priming. *Cognition*, 70(1).
- Nelson, H.E., (1991). National Adult Reading Test, 2nd ed. NFER-Nelson, Windsor, UK.
- Oakhill, J.V. (1982). Constructive processes in skilled and less skilled comprehenders memory for sentences. *British Journal of Educational Psychology*, 73, 13-20.
- Oakhill, J. V. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology*, 54, 31-39.
- Oakhill, J. V. (1996). Mental models in children's text comprehension. In J. V. Oakhill & A. Garnham (Eds.), *Mental models in cognitive science: essays in honour of Phil Johnson-Laird*. Hove, East Sussex: Psychology Press.
- Oakhill, J.V., Hartt, J., & Samols, D. (1995, April). Comprehension monitoring and working memory in good and poor comprehenders. Paper presented at the Experimental Psychology Society meeting, Cambridge, England. As cited in Cornoldi, C. & Oakhill, J.V. (Eds)(1996). *Reading Comprehension Difficulties: Processes and Intervention*. LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS 1996 Mahwah, New Jersey
- Oakhill, J. & Cain, K. (2000). Children's difficulties in text comprehension: Assessing Causal Issues. *Journal of Deaf Studies and Deaf Education*. 5:1. 51-59.
- O'Brien, E.J. & Albrecht, J.E., (1992). Comprehension strategies in the development of a mental model. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 777-784.
- Odegard, T.N, Holliday, R.E., Brainerd, C.J., & Reyna, V.F. (2008). Attention to global gist processing eliminates age effects in false memories. *Journal of Experimental Child Psychology*, 99(2), 96-113.
- Perfetti (1985). *Reading Ability*. New York, NY: Oxford.

- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading* 11(4), 357-383.
- Perfetti, C.A. & Hart, L. (2002). The Lexical Quality Hypothesis. In Verhoeven, C., Elbro, C. & Reitsma, P. (Eds.) *Precursors of functional literacy*. (p. 189-213). Amsterdam/Philadelphia: John Benjamins.
- Perfetti, C., Yang, C., & Schmalhofer, F. (2008). Comprehension Skill and Word-to-Text Integration Processes. *Applied Cognitive Psychology*, 22, 303–318.
- Reyna, V. F. & Kiernan, B. (1994). The development of gist versus verbatim memory in sentence recognition: effects of lexical familiarity, semantic content, encoding instructions and retention interval. *Developmental Psychology*, 30, 178-191.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: remembering words not presented on lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 803–814.
- Roediger, H. L., & McDermott, K. B. (1999). False alarms and false memories. *Psychological Review*, 106, 406–410.
- Sesma, H. W., Mahone, M.E., Levine, T., Eason, S.H., Cutting, L.E. (2009). The Contribution of Executive Skills to Reading Comprehension. *Child Neuropsychology* 15(3), 232-246
- Snodgrass, J.G. & Corwin, (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34-50.
- Spilich, G. J., Vesonder, G.T., Chiesi, H.L., & Voss, J.F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, 18(3), 257-380.
- Stadler, M.A., Roediger, H.L., & McDermott, K.B. (1999). Norms for words that create false memories. *Memory and Cognition*, 27, 494-500.
- Stanislaw, H. & Todorov, N. (1999). Calculation of signal detection theory measures, *Behaviour Research Methods, Instruments, & Computers*, 31, 137–149.
- Stanovich, K.E. & Cunningham, A.E. (1992). Studying the Consequences of Literacy Within a Literate Society – The Cognitive Correlates of Print Exposure. *Memory & Cognition*, 20(1), 51-68.

Stanovich, K.E., & West, R.F. (1989). Exposure to Print and Orthographic Processing. *Reading Research Quarterly*, 24(4), 402-433.

Trabasso, T. and P. van den Broek (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24(5), 612-630

Turner, M.L., & Engle, R.W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.

Unsworth, N., Redick, T.S., Heitz, R.P., Broadway, J.M., & Engle, R.W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, 17(6), 635 - 654.

Wagoner, S. A. (1983). Comprehension Monitoring: What It Is and What We Know about It. *Reading Research Quarterly* 18(3), 328-346.

Warburton, E., Wise, R.J.S., Price, C.J., Weiller, C., Harad, U., Ramsay, S., & Frackowiak, R.S.J. (1996). Noun and verb retrieval by normal subjects studies with PET. *Brain*, 119, 159-179.

Waters, G. S. & Caplan, D. (1996). The capacity theory of sentence comprehension: critique of Just and Carpenter (1992). *Psychological Review*, 103(4), 761-72.

Weekes, B.S., Hamilton, S., Oakhill, J., & Holliday, R.E. (2008). False recollection in children with reading comprehension difficulties. *Cognition*, 106(1), 222-233.

West, R.F. & Stanovich, K.E. (1991). The Incidental Acquisition of Information From Reading. *Psychological Science*, 2(5), 325-330.

Westbury, C. Buchanan, L. & Brown, N.R. (2002). Sounds of the neighbourhood: false memories and the structure of the phonological lexicon. *Journal of Memory and Language*, 40, 83-108

Whitney, P., Arnett, P.A., Driver, A., Budd, D. (2001). Measuring central executive functioning: What's in a reading span? *Brain and Cognition* 45(1), 1-14.

Yuill, N.M. and Oakhill, J.V. (1991) *Children's Problems in Text Comprehension: An Experimental Investigation*. Cambridge University Press.

Appendix 1: Materials Used in Chapter 3 (Experiment 1)

Study Lists and Recognition Tests for the Semantic and Phonological Versions of the DRM Paradigm (from Weekes et al., 2008)

Semantic study lists

Study List 1:

door glass pane shade ledge sill house open curtain frame

Study List 2:

bed rest awake tired dream awake snooze blanket doze slumber

Study List 3:

nose breathe sniff aroma hear see nostril whiff scent reek

Study List 4:

nurse sick lawyer medicine health hospital dentist physician ill patient

Study List 5:

sour candy sugar bitter good taste tooth nice honey soda

Study List 6:

table sit legs seat couch desk recliner sofa wood cushion

Semantic Recognition Test

(Related and unrelated distracters in *italics*, critical distracters in **bold**)

door nose *view nap* radio pane *stool* **sweet** *stench* sing bed sour **window** awake health scent tooth
snooze *breeze chocolate stethoscope snore* dentist curtain dress soldier legs hear *swivel* sniff couch
fragrance **sleep** handle dream *office heart* mug nurse war **doctor** house **smell** toe kick **chair** honey
lawyer dance wood crayon table sugar pencil

Phonological Study Lists

Study List 1:

chase keen wail wing cop wan weep watt lag wheat

Study List 2:

ring leave league tad lead leaf soot leash fun lean

Study List 3:

stole pole lathe bowl doze bake hole mole cane dole

Study List 4:

beach chip laugh bush doll bite bath bull bell kiss

Study List 5:

lone date file shown bone gauge hone phone soup tone

Study List 6:

rat site rash wrath bomb rag join cube rack ram

Phonological Recognition Test

(Related and unrelated distracters in *italics*, critical distracters in **bold**)

weep rat leave beach fun *sewn* bowl wing shown **leak** wash *leach* **wrap** cone wipe bag hot **will**
 bake keen chase leaf ring site *leap* fog zoom pole *boat* lathe doze *coal* toss *soul* mash **roll** bush
 laugh chip *beat* wag *rang* fool knife loan date file *game* hawk zone rash bomb *ran* soar

A' Equation for Recognition Performance in Experiment 1

A' is the nonparametric counterpart of d' (Snodgrass & Corwin, 1988) and is calculated using one of two different formulae depending on the relationship between hit and false alarm rates. For Experiment 1, the single formula suggested by Stanislaw and Todorov (1999) was used:

$$A' = 0.5 + \left[\text{sign}(H - F) \frac{(H - F)^2 + |H - F|}{4 \max(H, F) - 4HF} \right]$$

Appendix 2: Materials Used in Chapters 3, 4, and 5 (Experiments 1, 2, 3, and 4)

The Sentence Span Task (from McCallum, 2002) (used in Experiments 1, 2, 3, and 4)

(N.B. 'TA' = True Animate; 'FA' = False Animate; 'TI' = True Inanimate; 'FI' = False Inanimate)

Instructions:

Working Memory

Standard Sentence Span

"This is a memory test.

I will read out groups of sentences and after each one you have to say whether or not the sentence makes sense (Yes/No)

You also have to remember the last word of each sentence as you are making the judgements. When I get to the end of a group of sentences I will say "end" and you then have to list the words you have been remembering.

Try to recall them in the right order, but if you can't just recall as many words as possible

Do you understand? Let's try some examples"

[Use stimuli from 2 sentence groups (the subjects I tested were at ceiling on these), unless you believe this span size may be difficult for your subjects in which case you'll have to devise practice stimuli.]

Test:

2 [can be used as practice]

- | | |
|----|--|
| TA | When a boy grows up he is called a man |
| FA | There are many people who will be made by the bee |
| TI | Some mornings it is hard to get out of bed |
| FI | To mend your sock you must have indoor plants |
| FA | If you need a car you should cook the goat |
| TI | It is hard to sleep if you have just had a coffee |

TI A car will start more easily if kept in a **garage**
 FA When writing a book you will need a **cow**

FA A kitchen stool is usually kept in a **pig**
 FA You can sell all your jumpers to a **baker**

3

FA The largest insect that sings is the **flea**
 TA When you are ill it is best to go to a **doctor**
 FI When you are thirsty it is wise to have a **bath**

FA There are some who grow up to hunt the **florist**
 FI If you want to make toast you will be kept in a **shed**
 FI It is good exercise to run up and down a **vase**

TI To remove your sock you must take off your **shoe**
 TA You don't need to buy eggs if you own a **chicken**
 FI Gardening tools usually need some **bread**

TI When making a cup of tea you need a **kettle**
 FI The cook put pepper and salt in the lost **cushion**
 FI It is easy to open the door with a **coffee**

TI You can go to the library to borrow a **book**
 TA If you need meat you should go to the **butcher**
 TA There are many different breeds of **dog**

4

FA You will often fall asleep if you see a **ladybird**
 TA Sometimes naughty children will be rude to their **teacher**
 FI If you go to the library you must wear a **coat**
 FI Candles and pictures may jump up onto the **shelf**

FA The pet shop keeps many different breeds of **grocer**
 TA One of the prettiest insects is the **butterfly**
 TA The policeman will give chase if he sees a **thief**
 FI We can eat the honey that is singing to the **carpet**

TA The elephant is said to be scared of the **mouse**
 TI If you want to make toast you will need some **toast**
 FA The kitchen table will often make a living as an **actor**
 FI A car will start more easily when lying on the **sofa**

TI At the end of a long day it is nice to have a **sleep**
 FA You need to like to go skiing to be a **gerbil**
 FI Many people are allergic to their own **tie**
 TA There are some who like to hunt the **fox**

FI One of the prettiest insects is the **table**
 TA You must believe in God if you are a **bishop**
 TI If it is cold outside you should wear a **jumper**
 FA It is easier to sing if you have just eaten a **rabbit**

5

TA There are many people who will run away from a **spider**
 FI If it is cold outside you should go in a **fridge**
 FA You must keep bicycles chained up near a **hamster**
 TI Gardening tools are usually kept in a **shed**
 TA An example of a parasitic insect is the **flea**

TA It is irritating to be bitten by a **mosquito**
 TI Cut flowers should be put into water in a **vase**
 TI Before baking it is wise to heat up the **oven**
 TI In the summer the garden is full of **flowers**
 FI Food that must be kept cold should wear a **sock**

FA A teacher will often give medicine to a **doctor**
 TI It is good exercise to run up and down the **stairs**
 TA One creature that is famed for climbing is the **goat**
 FA If you believe in God you will become a **pilot**
 FA When you are ill it is best to be rude to the **donkey**

TA In the summer you will often see a **caterpillar**
 FA Before baking it is wise to be afraid of a **horse**
 FA You need years of training to give food to a **mouse**
 TI If you go out in the rain it is wise to wear a **coat**
 FA It is bad luck to put your shoes on the **butterfly**

FI Some mornings an insect is hard to get out of **bed**
 TI You can't get inside if you have lost your **key**
 TI A cook will often put pepper and salt in the **soup**
 TA You can feed all your leftovers to a **pig**
 FI In the summer you will often be kept in a **garage**

6

TA You need to like cooking to be a **baker**
 TI If you don't wipe your feet you may ruin the **carpet**
 TA Many young children like to keep a **pet**
 FI If you go out in the rain it is wise to borrow a **book**
 TI It is bad luck to put your shoes on the **table**
 FA You need eggs if you want to be a **bishop**

TA You need years of training to become a **surgeon**
 TI Ornaments and photos can be kept on the **shelf**
 FA One of the rarest wild animals is the **teacher**
 TI It is easy to fall asleep when lying on the **sofa**
 FA Many young kittens will shake hands with a **man**
 FI Cut flowers should be put into water in an **eggcup**

TA It is hard to make a living as an **actor**
 FI It is irritating to heat up the **oven**
 FA Sometimes the unruly children would drink the **thief**
 TA We can eat the honey that is made by the **bee**
 FA It is nice to be sung to by a **mosquito**
 FI If you don't wipe your feet you may need a **kettle**

TA There are many occasions when you need a **florist**
 FI An uncomfortable chair is inside with the **vicar**
 TI Many schools and clubs have their own **tie**
 FA When a boy eats meat he is called a **fox**
 TI An uncomfortable chair is improved with a **cushion**
 FI In the summer the garden must take off your **shoe**

FI You can't get improved if you have lost your **soup**
 TA You must keep dogs under control when near **sheep**
 TI Food that must be kept cold should go in a **fridge**
 FI At the end of the day you should eat **flowers**
 TA Jesus was said to have ridden on a **donkey**

TI For the house there are many types of indoor **plants**

Word-level Tasks

Pseudohomophone/3 minutes (used in Experiments 1, 2, 3, and 4)

Instructions:

Name:

On the next page, you will see a list of 100 letter strings. Some of the strings in the list, if you sound them out, sound like real words (for example ‘gote’ sounds like ‘goat’) – even if they are not spelled like real words. Some of the strings, when you sound them out, do not sound like real words. Please read the strings and check each one that sounds like a real word. Do not guess, but only check those you know to be words. Remember, some of them are made-up words, so guessing can be easily detected.

You have 3 minutes to do this task.

Test:

| | | | | |
|------------------------------------|-------------------------------------|--|---------------------------------------|------------------------------------|
| <input type="checkbox"/> kake | <input type="checkbox"/> deace | <input type="checkbox"/> gairubteigh | <input type="checkbox"/> warld | <input type="checkbox"/> fite |
| <input type="checkbox"/> thair | <input type="checkbox"/> oquird | <input type="checkbox"/> keitreeot | <input type="checkbox"/> jope | <input type="checkbox"/> feefail |
| <input type="checkbox"/> fether | <input type="checkbox"/> mijen | <input type="checkbox"/> thrue | <input type="checkbox"/> joak | <input type="checkbox"/> thurd |
| <input type="checkbox"/> neinton | <input type="checkbox"/> dak | <input type="checkbox"/> grafight | <input type="checkbox"/> fither | <input type="checkbox"/> tirn |
| <input type="checkbox"/> colesture | <input type="checkbox"/> skwolb | <input type="checkbox"/> hoate | <input type="checkbox"/> kwartet | <input type="checkbox"/> eer |
| <input type="checkbox"/> theer | <input type="checkbox"/> lote | <input type="checkbox"/> kelone | <input type="checkbox"/> teech | <input type="checkbox"/> widome |
| <input type="checkbox"/> derty | <input type="checkbox"/> plice | <input type="checkbox"/> jimpozyum | <input type="checkbox"/> muroseize | <input type="checkbox"/> urthkweik |
| <input type="checkbox"/> broave | <input type="checkbox"/> flote | <input type="checkbox"/> strate | <input type="checkbox"/> shart | <input type="checkbox"/> timnastix |
| <input type="checkbox"/> nerr | <input type="checkbox"/> seaf | <input type="checkbox"/> amiture | <input type="checkbox"/> anteik | <input type="checkbox"/> roak |
| <input type="checkbox"/> ploor | <input type="checkbox"/> cliss | <input type="checkbox"/> ulsir | <input type="checkbox"/> heitr | <input type="checkbox"/> onsaumbel |
| <input type="checkbox"/> turt | <input type="checkbox"/> puteighter | <input type="checkbox"/> vazhyoual | <input type="checkbox"/> trane | <input type="checkbox"/> eckseed |
| <input type="checkbox"/> bloe | <input type="checkbox"/> shaip | <input type="checkbox"/> speschelst | <input type="checkbox"/> trator | <input type="checkbox"/> horl |
| <input type="checkbox"/> craul | <input type="checkbox"/> edepust | <input type="checkbox"/> shatcheure | <input type="checkbox"/> phrantcheyes | <input type="checkbox"/> caim |
| <input type="checkbox"/> quaynt | <input type="checkbox"/> filce | <input type="checkbox"/> fearce | <input type="checkbox"/> baik | <input type="checkbox"/> crail |
| <input type="checkbox"/> dofter | <input type="checkbox"/> eap | <input type="checkbox"/> phoarbearense | <input type="checkbox"/> chonseed | <input type="checkbox"/> gaim |
| <input type="checkbox"/> pebl | <input type="checkbox"/> threp | <input type="checkbox"/> aknoleque | <input type="checkbox"/> matcheune | <input type="checkbox"/> fairce |
| <input type="checkbox"/> saip | <input type="checkbox"/> saif | <input type="checkbox"/> skwyde | <input type="checkbox"/> beal | <input type="checkbox"/> bote |
| <input type="checkbox"/> baionette | <input type="checkbox"/> skweez | <input type="checkbox"/> reash | <input type="checkbox"/> serkyouler | <input type="checkbox"/> thord |
| <input type="checkbox"/> jadwarr | <input type="checkbox"/> ohooodeoh | <input type="checkbox"/> ferst | <input type="checkbox"/> bape | <input type="checkbox"/> viknaime |
| <input type="checkbox"/> world | <input type="checkbox"/> traif | <input type="checkbox"/> youroveu | <input type="checkbox"/> indiseeze | <input type="checkbox"/> poantil |

Pseudohomophone Circle (used in Experiments 1, 2, 3, and 4)**Instructions**

NAME:

DATE:

NONSENSE WORDS THAT SOUND LIKE REAL WORDS

In this task you will see a series of rows, each with three nonsense words. None of these are real words. However, one of them would sound like a real word if you sounded it out in your head. Your task is to indicate which one would sound like a real word by circling it.

We will start with five practice items, and I will tell you the correct answers to those items and ask if you have any questions before we start the main test.

PRACTICE ITEMS

- | | | |
|---------|------|-------|
| 1. nite | kile | hote |
| 2. beal | bair | rabe |
| 3. ean | eer | eap |
| 4. glew | plue | glay |
| 5. slod | stum | steem |

Now you are ready to start the main test. You should try to work accurately, but also quite quickly - otherwise you might not have time to finish.

****DO NOT TURN OVER**

UNTIL YOU ARE ASKED TO DO SO**

TEST ITEMS

| | | |
|--------|--------|--------|
| gan | pab | kat |
| poal | hoil | mool |
| kape | dape | lape |
| pid | lud | sed |
| breen | bloan | clain |
| gass | hask | wask |
| crish | wosh | hesh |
| coam | baim | goam |
| suv | wiv | hav |
| fard | werd | lurd |
| fude | fode | gade |
| doin | stoon | boan |
| sharf | slout | skore |
| slin | scip | shill |
| coff | thriff | pluff |
| gizz | duzz | tazz |
| poost | moast | froust |
| chooze | goote | mooze |
| caim | goim | neem |
| blad | stid | flud |
| stune | pline | trane |
| lite | fote | nate |
| cleen | bain | vown |
| horge | horce | horve |
| droom | creem | staim |

| | | |
|---------|-----------|----------|
| skait | plout | steet |
| fleek | brouk | shaik |
| shoop | crawp | sleap |
| pellar | senter | nertain |
| strook | squaik | speek |
| dysical | fotograph | barmacy |
| teer | gair | cour |
| cree | spoo | flie |
| roke | noze | clobe |
| sirst | kirch | bight |
| cloor | fleer | scair |
| dircus | migar | sircle |
| phan | shog | chun |
| stee | floo | blai |
| firch | sirst | sicks |
| nule | rale | jile |
| shurt | spart | dort |
| tooch | reech | paich |
| fout | lait | doot |
| fone | phote | toaf |
| smeak | joak | paik |
| phinny | hanny | munny |
| flate | gite | bote |
| sentury | bertainly | mertify |
| naim | soom | coim |
| peneral | pentle | jenerous |

** TURN TO NEXT PAGE**

skwash shreeze skrad

Orthographic Transformation Task (used in Experiments 1 and 2)

This is a test of your ability to think about how words sound, regardless of how they're spelled. Each question has a word. Say this word silently to yourself. Ex. CRUISE (say to yourself, /krooz/).

Then look at the sound to be removed from the word. Say the word silently to yourself, without the sound. Write the word that you said. Ex. Remove the /z/ sound. (say to yourself, /kroo/. Write the word crew.)

Then look at the sound to be replaced in the spot where you removed the first sound. Say the word silently to yourself, adding this sound. Write the word that you said. Ex. Add the /l/ sound. (say to yourself, /krool/. Write the word cruel.)

Here's another example.

SPEAK remove the /p/ sound seek add the /l/ sound sleek

Sometimes you might actually have to split apart the sounds made by one letter.

MIXED remove the /k/ sound missed add the /d/ sound midst

Sometimes you might have to put together sounds made by several letters.

PHONICS remove the /f/ sound onyx add the /t/ sound tonics

*****Hint, Hint: Everything you write should be a real word, correctly spelled.*****

Now complete the following items:

- | | | | | |
|-------------|----------------------------|--------------|--------------------|----------------|
| 1. MIDDLE | remove the /d/ sound | <u>mill</u> | add the /s/ sound | <u>missile</u> |
| 2. QUEEN | remove the /w/ sound | <u>keen</u> | add the /l/ sound | <u>clean</u> |
| 3. NICKEL | remove the /k/ sound | <u>nil</u> | add the /b/ sound | <u>nibble</u> |
| 4. QUAKE | remove the first /k/ sound | <u>wake</u> | add the /uh/ sound | <u>awake</u> |
| 5. MOTION | remove the /m/ sound | <u>ocean</u> | add the /l/ sound | <u>lotion</u> |
| 6. WRAPPED | remove the /r/ sound | <u>apt</u> | add the /t/ sound | <u>tapped</u> |
| 7. CAUGHT | remove the /k/ sound | <u>ought</u> | add the /t/ sound | <u>taught</u> |
| 8. PAGE | remove the /j/ sound | <u>pay</u> | add the /n/ sound | <u>pain</u> |
| 9. LAUGHTER | remove the /l/ sound | <u>after</u> | add the /r/ sound | <u>rafter</u> |
| 10. SKY | remove the /k/ sound | <u>sigh</u> | add the /p/ sound | <u>spy</u> |
| 11. FARCE | remove the /s/ sound | <u>far</u> | add the /m/ sound | <u>farm</u> |
| 12. COLONEL | remove the /n/ sound | <u>curl</u> | add the /d/ sound | <u>curdle</u> |
| 13. RACKS | remove the /r/ sound | <u>axe</u> | add the /t/ sound | <u>tax</u> |
| 14. MIGHT | remove the /t/ sound | <u>my</u> | add the /s/ sound | <u>mice</u> |
| 15. SPOT | remove the /p/ sound | <u>sot</u> | add the /l/ sound | <u>slot</u> |
-

Appendix 3: Materials Used in Chapter 5 (Experiments 3 and 4)

Coherence and Cohesion Paradigm (from Ferstl, 2002) (used in Experiments 3 and 4)

Instructions:

"This is an experiment about reading.

You will see two sentences, one after the other. They are short stories: some make sense, some don't.

As soon as you have finished reading each sentence, press 'NEXT' to carry on.

After you have pressed the 'NEXT' key for a second time, you will see a question mark (?).

If the two sentences make sense together press 'YES'. If the two sentences do not make sense together press 'NO'.

Please respond as quickly as possible. Press 'NEXT' to begin."

Target and Context Sentences (version 1 of 4):

Coherent/Cohesive

A big anti-war demonstration has been announced.
Therefore the whole town is swarming with police.

Barbara leaves the lecture hall after her exam completely exhausted.
Her friends are waiting at the door with a bottle of champagne.

In the city there are many offices available.
Therefore the rents are relatively low.

John doesn't recognise who is waving to him in the park.
His glasses are lying on his desk at home.

Last night there was a terrible storm.
Today the roads are covered in branches.

Maria is a single mother.
Thus every penny counts.

Marjorie has just learned to swim.
Her mother is standing anxiously near the deep end of the pool.

Martin's thoughts drifted off for a moment.
Because of this there was nearly a collision.

Matt buys a bouquet of flowers.
His colleague has given birth to a baby girl.

Mr. Gardner has a bad cold.
In his flat handkerchiefs are lying around everywhere.

Mrs. Jones has a hair cut.
Tomorrow will be her staff outing.

Outside it is cold and wet.
This is why Max has to take off his glasses when entering the flat.

Rob has booked tickets for the opera.
Thus Julia rings her baby sitter.

The account was overdrawn by a lot.
Therefore the automatic cash dispenser swallowed the card.

The bills have not been paid for a long time.
Now the phone company will disconnect the phone line.

The computer room has security camera surveillance.
This should prevent theft.

The flight has been cancelled at short notice.
Now Toby is sitting in the departure lounge swearing to himself.

The fountain pen has stopped working.
Its cartridge is presumably empty.

The new glue is not very good.
The wallpaper came off the wall straight away.

The new government has increased the tax on energy.
Now petrol prices will go up.

The radio is rarely switched on.
Its volume control is no longer working.

The surgery involved much blood loss.
Afterwards a lot of gowns had to be cleaned.

The travel news announced long traffic jams on the motorway.
Therefore Fred left an hour earlier.

The X-ray picture is rather blurred.
The tooth looks enormous on it.

Coherent/Incohesive

A new supermarket has been opened.
The income of the local bakery has gone down.

All advance tickets were sold out after a few days.
There were long queues at the box office.

In Scotland there are lots of midges.
Simone has armed herself with insect repellent.

It has been raining for days.
All the paths have turned into mud.

Maggie has placed an ad in the paper.
The telephone is ringing constantly.

Marilyn has problems with her son.
The teacher recommends a book about modern parenting.

Matthew has just learned to roller skate.
The knees are grazed all the time.

Mr. Simmons takes his newspaper to work.
The 45-minute train trip can be rather boring.

On Guy Fawkes' night a rocket exploded too early.
The village hall burned down.

Robert heard on the weather forecast that there would be black ice on the roads.
Luckily, the season ticket is still valid.

The ash trays were not emptied last night.
The office smells like a pub.

The bloated feeling after a meal is very unpleasant.
Martin carries digestive tablets.

The campers want to go to bed.
Water is being poured on the fire.

The car is speeding round the corner.
The tyres are squealing loudly.

The car was left on the driveway for a long time.
The windows have to be de-iced.

The danger of frost is over.
Mrs. Smith plants her window boxes.

The electricity has failed in the shop on the corner.
Frozen pizza is on special offer.

The heating broke down over the weekend.
It was bitterly cold in the school.

The postal service was on strike.
Not a single Christmas parcel reached its recipient on time.

The summer was very hot.
Suntan lotion was nearly sold out.

The window is not quite shut.
The curtains are fluttering constantly.

Toby tidies up his flat.

Visitors are expected the following day.

Tomorrow is the day when the rubbish is collected.
Tom puts his dustbin outside.

When Monica opened her door a pile of mail lay on the mat.
Nobody had forgotten the birthday.

Incoherent/Cohesive

A mistake occurred during the roadworks.
Therefore all the textbooks are out on loan.

A new detergent has come on the market.
Since then it has become much quieter in the flat.

An hour has passed.
Consequently the car will not start tomorrow morning.

Anne left for Paris three months ago.
She breaks out in a cold sweat.

Annette is going to wallpaper her flat.
Her heart starts beating faster every time the phone rings.

Carmen is giving a lecture.
Her back and thighs are completely red.

Icy rain made the streets dangerous.
They had to be assembled again in their proper order.

In January Mrs. Crow is looking for her sun tan lotion.
Her friend is celebrating his 21st birthday.

Kim fell from the ski lift.
At that moment on the ground floor a window pane crashes.

Lisa flies to the Caribbean for her holiday.
Her tray has fallen down twice already.

Monica has lost weight.

Then a strong coffee is essential.

Next week is the company dinner dance.
Therefore Marjorie turns down her stereo.

Nicole would like to return to her job after her maternity leave.
She puts fresh towels into the bathroom.

The computer has crashed.
Its doors swing open by themselves.

The cup final is being shown on TV.
Consequently the sea level will rise.

The road leading to the fun fair is closed to cars.
For this, the use of a medicinal mouthwash is recommended.

The shower curtain did not quite shut.
In that region several houses collapsed.

The university holidays are in August.
Therefore the glasses often rattle in the cupboard.

There were roadworks near the motorway junction.
Thus more language schools are offering Slavic languages.

Two traffic wardens walk past.
Next week is their wedding anniversary.

Vera cannot concentrate on her work.
That is why pink baby clothes are bought.

Walking long distances when sightseeing can lead to blisters.
Therefore the insurance company will increase the premium.

Wool can often shrink when put into the washing machine.
Hence smoking in filling stations is strictly prohibited.

Yesterday there was a big family get together.
Therefore Elly has her bike light checked.

Incoherent/Incohesive

Before the lecture the transparencies fell on the floor.
All the buses were delayed.

Cassie cannot concentrate on her work today.
All the colleagues are envious.

Eric has to do a night shift at the hospital.
The old trousers fit again.

Frank hides his purchase from his wife.
Peter jumps up and parks his car somewhere else.

Global warming will cause glaciers to melt.
The streets are completely empty.

In autumn the days are getting shorter.
Mrs. McCarthy is pleased to have a dish washer.

It is nearly 10 pm.
Mrs. Robinson takes her best dress to be dry cleaned.

Jenny buys a bottle of champagne.
The holiday is skiing in the mountains.

Judith's heart starts pounding when she gives a talk.
Thomas has a high telephone bill.

Many heavy lorries drive past the house.
The campus is almost deserted.

Martine went to sleep on the beach.
The hands are trembling.

Mrs. Fine hopes to have a little girl.
A short walk gives new energy.

Mrs. Ford is having visitors.
The offer of a computing course sounds interesting.

Petrol fumes are highly inflammable.
The thick pullover should be washed by hand.

Sally is waiting to hear from Michael whom she fancies.
The old pullover and working trousers are already laid out.

The chemistry exam will take place soon.
The telephones are dead in the whole village.

The children are playing football on the patio.
Both legs are in plaster.

The cupboard is tilted.
The monitor is only displaying meaningless symbols.

The damage from the accident was greater than was anticipated.
A travel bag should always contain plasters.

The epicentre of the earthquake was near Los Angeles.
The bathroom floor is rather wet.

The headlights are still on.
The cake should be ready.

The markets of the future lie in Eastern Europe.
The traffic jams extended for several miles.

The railway line at the back of the house has been shut down.
The towels are soft and smell fresh.

The toothbrush cannot easily get into the gaps between the teeth.
A free bus service is provided.

Target and Context Sentences (version 2 of 4):

Coherent/Cohesive

Before the lecture the transparencies fell on the floor.
They had to be assembled again in their proper order.

Cassie cannot concentrate on her work today.

Her tray has fallen down twice already.

Eric has to do a night shift at the hospital.
Then a strong coffee is essential.

Frank hides his purchase from his wife.
Next week is their wedding anniversary.

Global warming will cause glaciers to melt.
Consequently the sea level will rise.

In Scotland there are lots of midges.
Therefore Simone has armed herself with insect repellent.

In the autumn the days are getting shorter.
Therefore Elly has her bike light checked.

It is nearly 10 pm.
Therefore Marjorie turns down her stereo.

Jenny buys a bottle of champagne.
Her friend is celebrating his 21st birthday.

Judith's heart starts pounding when she gives a talk.
She breaks out in a cold sweat.

Many heavy lorries drive past the house.
Therefore the glasses often rattle in the cupboard.

Martine went to sleep on the beach.
Her back and thighs are completely red.

Mrs. Fine hopes to have a little girl.
That is why pink baby clothes are bought.

Mrs. Ford is having visitors.
She puts fresh towels into the bathroom.

Petrol fumes are highly inflammable.
Hence smoking in filling stations is strictly prohibited.

Sally is waiting to hear from Michael whom she fancies.
Her heart starts beating faster every time the phone rings.

The chemistry exam will take place soon.
Therefore all the textbooks are out on loan.

The children are playing football on the patio.
At that moment on the ground floor a window pane crashes.

The cupboard is tilted.
Its doors swing open by themselves.

The damage from the accident was greater than was anticipated.
Therefore the insurance company will increase the premium.

The epicentre of the earthquake was near Los Angeles.
In that region several houses collapsed.

The headlights are still on.
Consequently the car will not start tomorrow morning.

The railway line at the back of the house has been shut down.
Since then it has become much quieter in the flat.

Toothbrushes cannot easily get into the gaps between the teeth.
For this, the use of a medicinal mouthwash is recommended.

Coherent/Incohesive

A mistake occurred during the roadworks.
The telephones are dead in the whole village.

A new detergent has come on the market.
The towels are soft and smell fresh.

An hour has passed.
The cake should be ready.

Anne left for Paris three months ago.
Thomas has a high telephone bill.

Annette is going to wallpaper her flat.
The old pullover and working trousers are already laid out.

Carmen is giving a lecture.
The hands are trembling.

Icy rain made the streets dangerous.
All the buses were delayed.

In January Mrs. Crow is looking for her sun tan lotion.
The holiday is skiing in the mountains.

Kim fell from the ski lift.
Both legs are in plaster.

Lisa flies to the Caribbean for her holiday.
All the colleagues are envious.

Monica has lost weight.
The old trousers fit again.

Next week is the company dinner dance.
Mrs. Robinson takes her best dress to be dry cleaned.

Nicole would like to return to her job after her maternity leave.
The offer of a computing course sounds interesting.

The computer has crashed.
The monitor is only displaying meaningless symbols.

The cup final is being shown on TV.
The streets are completely empty.

The road leading to the fun fair is closed to cars.
A free bus service is provided.

The shower curtain did not quite shut.
The bathroom floor is rather wet.

The travel news announced long traffic jams on the motorway.

Fred left an hour earlier.

The university holidays are in August.
The campus is almost deserted.

Two traffic wardens walk past.
Peter jumps up and parks his car somewhere else.

Vera cannot concentrate on her work.
A short walk gives new energy.

Walking long distances when sightseeing can lead to blisters.
A travel bag should always contain plasters.

Wool can often shrink when put into the washing machine.
The thick pullover should be washed by hand.

Yesterday there was a big family get-together.
Mrs. McCarthy is pleased to have a dish washer.

Incoherent/Cohesive

A new supermarket has been opened.
Afterwards a lot of gowns had to be cleaned.

All advance tickets were sold out after a few days.
Now petrol prices will go up.

It has been raining for days.
Now the phone company will disconnect the phone line.

Maggie has placed an ad in the paper.
Thus Julia rings her baby sitter.

Marilyn has problems with her son.
Tomorrow will be her staff outing.

Matt buys a bouquet of flowers.
Otherwise the 45-minute train trip can be rather boring.

Matthew has just learned to roller skate.
In his flat handkerchiefs are lying around everywhere.

On Guy Fawkes' night a rocket exploded too early.
Therefore the automatic cash dispenser swallowed the card.

Robert heard on the weather forecast that there would be black ice on the roads.
His glasses are lying on his desk at home.

The ash trays were not emptied last night.
Today the roads are covered in branches.

The bloated feeling after a meal is very unpleasant.
This is why Max has to take off his glasses when entering the flat.

The campers want to go to bed.
Thus every penny counts.

The car is speeding round the corner.
The tooth looks enormous on it.

The car was left on the driveway for a long time.
The wallpaper came off the wall straight away.

The danger of frost is over.
Her friends are waiting at the door with a bottle of champagne.

The electricity has failed in the shop on the corner.
This should prevent theft.

The heating broke down over the weekend.
Therefore the whole town is swarming with police.

The markets of the future lie in Eastern Europe.
From that point the traffic jams extended for several miles.

The postal service was on strike.
Therefore the rents are relatively low.

The summer was very hot.
Its cartridge is presumably empty.

The window is not quite shut.
Its volume control is no longer working.

Toby tidies up his flat.
Because of this there was nearly a collision.

Tomorrow is the day when the garbage is collected.
Now Toby is sitting in the departure lounge swearing to himself.

When Monica opened her door a pile of mail lay on the mat.
Her mother is standing anxiously near the deep end of the pool.

Incoherent/Incohesive

Barbara leaves the lecture hall after her exam completely exhausted.
Mrs. Smith plants her window boxes.

A big anti-war demonstration has been announced.
It was bitterly cold in the school.

In the city there are many offices available.
Not a single Christmas parcel reached its recipient on time.

John doesn't recognise who is waving to him in the park.
Luckily, the season ticket is still valid.

Last night there was a terrible storm.
The office smells like a pub.

Maria is a single mother.
Water is being poured on the fire.

Marjorie has just learned to swim.
Nobody had forgotten the birthday.

Martin's thoughts drifted off for a moment.
Visitors are expected the following day.

Mr. Gardner has a bad cold.

The knees are grazed all the time.

Mr. Simmons takes his newspaper to work.
A colleague has given birth to a baby girl.

Mrs. Jones has a hair cut.
The teacher recommends a book about modern parenting.

Outside it is cold and wet.
Martin carries digestive tablets.

Rob has booked tickets for the opera.
The telephone is ringing constantly.

The account was overdrawn by a lot.
The village hall burned down.

The bills have not been paid for a long time.
All the paths have turned into mud.

The computer room has security camera surveillance.
Frozen pizza is on special offer.

The flight has been cancelled at short notice.
Tom puts his dustbin outside.

The fountain pen has stopped working.
Suntan lotion was nearly sold out.

The new glue is not very good.
The windows have to be de-iced.

The new government has increased the tax on energy.
There were long queues at the box office.

The radio is rarely switched on.
The curtains are fluttering constantly.

The surgery involved much blood loss.
The income of the local bakery has gone down.

The X-ray picture is rather blurred.
The tyres are squealing loudly.

There were roadworks near the motorway junction.
Many language schools are offering Slavic languages.

Target and Context Sentences (version 3 of 4):

Coherent/Cohesive

A big anti-war demonstration has been announced.
Therefore the whole town is swarming with police.

Barbara leaves the lecture hall after her exam completely exhausted.
Her friends are waiting at the door with a bottle of champagne.

In the city there are many offices available.
Therefore the rents are relatively low.

John doesn't recognise who is waving to him in the park.
His glasses are lying on his desk at home.

Last night there was a terrible storm.
Today the roads are covered in branches.

Maria is a single mother.
Thus every penny counts.

Marjorie has just learned to swim.
Her mother is standing anxiously near the deep end of the pool.

Martin's thoughts drifted off for a moment.
Because of this there was nearly a collision.

Matt buys a bouquet of flowers.
His colleague has given birth to a baby girl.

Mr. Gardner has a bad cold.
In his flat handkerchiefs are lying around everywhere.

Mrs. Jones has a hair cut.
Tomorrow will be her staff outing.

Outside it is cold and wet.
This is why Max has to take off his glasses when entering the flat.

Rob has booked tickets for the opera.
Thus Julia rings her baby sitter.

The account was overdrawn by a lot.
Therefore the automatic cash dispenser swallowed the card.

The bills have not been paid for a long time.
Now the phone company will disconnect the phone line.

The computer room has security camera surveillance.
This should prevent theft.

The flight has been cancelled at short notice.
Now Toby is sitting in the departure lounge swearing to himself.

The fountain pen has stopped working.
Its cartridge is presumably empty.

The new glue is not very good.
The wallpaper came off the wall straight away.

The new government has increased the tax on energy.
Now petrol prices will go up.

The radio is rarely switched on.
Its volume control is no longer working.

The surgery involved much blood loss.
Afterwards a lot of gowns had to be cleaned.

The travel news announced long traffic jams on the motorway.
Therefore Fred left an hour earlier.

The X-ray picture is rather blurred.
The tooth looks enormous on it.

Coherent/Incohesive

A new supermarket has been opened.
The income of the local bakery has gone down.

All advance tickets were sold out after a few days.
There were long queues at the box office.

In Scotland there are lots of midges.
Simone has armed herself with insect repellent.

It has been raining for days.
All the paths have turned into mud.

Maggie has placed an ad in the paper.
The telephone is ringing constantly.

Marilyn has problems with her son.
The teacher recommends a book about modern parenting.

Matthew has just learned to roller skate.
The knees are grazed all the time.

Mr. Simmons takes his newspaper to work.
The 45-minute train trip can be rather boring.

On Guy Fawkes' night a rocket exploded too early.
The village hall burned down.

Robert heard on the weather forecast that there would be black ice on the roads.
Luckily, the season ticket is still valid.

The ash trays were not emptied last night.
The office smells like a pub.

The bloated feeling after a meal is very unpleasant.
Martin carries digestive tablets.

The campers want to go to bed.
Water is being poured on the fire.

The car is speeding round the corner.
The tyres are squealing loudly.

The car was left on the driveway for a long time.
The windows have to be de-iced.

The danger of frost is over.
Mrs. Smith plants her window boxes.

The electricity has failed in the shop on the corner.
Frozen pizza is on special offer.

The heating broke down over the weekend.
It was bitterly cold in the school.

The postal service was on strike.
Not a single Christmas parcel reached its recipient on time.

The summer was very hot.
Suntan lotion was nearly sold out.

The window is not quite shut.
The curtains are fluttering constantly.

Toby tidies up his flat.
Visitors are expected the following day.

Tomorrow is the day when the rubbish is collected.
Tom puts his dustbin outside.

When Monica opened her door a pile of mail lay on the mat.
Nobody had forgotten the birthday.

Incoherent/Cohesive

A mistake occurred during the roadworks.
Therefore all the textbooks are out on loan.

A new detergent has come on the market.
Since then it has become much quieter in the flat.

An hour has passed.
Consequently the car will not start tomorrow morning.

Anne left for Paris three months ago.
She breaks out in a cold sweat.

Annette is going to wallpaper her flat.
Her heart starts beating faster every time the phone rings.

Carmen is giving a lecture.
Her back and thighs are completely red.

Icy rain made the streets dangerous.
They had to be assembled again in their proper order.

In January Mrs. Crow is looking for her sun tan lotion.
Her friend is celebrating his 21st birthday.

Kim fell from the ski lift.
At that moment on the ground floor a window pane crashes.

Lisa flies to the Caribbean for her holiday.
Her tray has fallen down twice already.

Monica has lost weight.
Then a strong coffee is essential.

Next week is the company dinner dance.
Therefore Marjorie turns down her stereo.

Nicole would like to return to her job after her maternity leave.
She puts fresh towels into the bathroom.

The computer has crashed.
Its doors swing open by themselves.

The cup final is being shown on TV.
Consequently the sea level will rise.

The road leading to the fun fair is closed to cars.

For this, the use of a medicinal mouthwash is recommended.

The shower curtain did not quite shut.
In that region several houses collapsed.

The university holidays are in August.
Therefore the glasses often rattle in the cupboard.

There were roadworks near the motorway junction.
Thus more language schools are offering Slavic languages.

Two traffic wardens walk past.
Next week is their wedding anniversary.

Vera cannot concentrate on her work.
That is why pink baby clothes are bought.

Walking long distances when sightseeing can lead to blisters.
Therefore the insurance company will increase the premium.

Wool can often shrink when put into the washing machine.
Hence smoking in filling stations is strictly prohibited.

Yesterday there was a big family get together.
Therefore Elly has her bike light checked.

Incoherent/Incohesive

Before the lecture the transparencies fell on the floor.
All the buses were delayed.

Cassie cannot concentrate on her work today.
All the colleagues are envious.

Eric has to do a night shift at the hospital.
The old trousers fit again.

Frank hides his purchase from his wife.
Peter jumps up and parks his car somewhere else.

Global warming will cause glaciers to melt.
The streets are completely empty.

In autumn the days are getting shorter.
Mrs. McCarthy is pleased to have a dish washer.

It is nearly 10 pm.
Mrs. Robinson takes her best dress to be dry cleaned.

Jenny buys a bottle of champagne.
The holiday is skiing in the mountains.

Judith's heart starts pounding when she gives a talk.
Thomas has a high telephone bill.

Many heavy lorries drive past the house.
The campus is almost deserted.

Martine went to sleep on the beach.
The hands are trembling.

Mrs. Fine hopes to have a little girl.
A short walk gives new energy.

Mrs. Ford is having visitors.
The offer of a computing course sounds interesting.

Petrol fumes are highly inflammable.
The thick pullover should be washed by hand.

Sally is waiting to hear from Michael whom she fancies.
The old pullover and working trousers are already laid out.

The chemistry exam will take place soon.
The telephones are dead in the whole village.

The children are playing football on the patio.
Both legs are in plaster.

The cupboard is tilted.
The monitor is only displaying meaningless symbols.

The damage from the accident was greater than was anticipated.
A travel bag should always contain plasters.

The epicentre of the earthquake was near Los Angeles.
The bathroom floor is rather wet.

The headlights are still on.
The cake should be ready.

The markets of the future lie in Eastern Europe.
The traffic jams extended for several miles.

The railway line at the back of the house has been shut down.
The towels are soft and smell fresh.

The toothbrush can not easily get into the gaps between the teeth.
A free bus service is provided.

Target and Context Sentences (version 4 of 4):

Coherent/Cohesive

A mistake occurred during the roadworks.
Now the telephones are dead in the whole village.

A new detergent has come on the market.
Since then the towels are soft and smell fresh.

An hour has passed.
Now the cake should be ready.

Icy rain made the streets dangerous.
Therefore all the buses were delayed.

In January Mrs. Crow is looking for her sun tan lotion.
Her holiday is skiing in the mountains.

Monica has lost weight.
Now her old trousers fit again.

Next week is the company dinner dance.

Thus Mrs. Robinson takes her best dress to be dry cleaned.

The computer has crashed.
Its monitor is only displaying meaningless symbols.

The cup final is being shown on TV.
Thus the streets are completely empty.

The danger of frost is over.
Now Mrs. Smith plants her window boxes.

The road leading to the fun fair is closed to cars.
However, a free bus service is provided.

The shower curtain did not quite shut.
Therefore the bathroom floor is rather wet.

The university holidays are in August.
During that month the campus is almost deserted.

There were roadworks near the motorway junction.
From that point the traffic jams extended for several miles.

Two traffic wardens walk past.
Immediately Peter jumps up and parks his car somewhere else.

Walking long distances when sightseeing can lead to blisters.
That is why a travel bag should always contain plasters.

Wool can often shrink when put into the washing machine.
Therefore the thick pullover should be washed by hand.

Yesterday there was a big family get-together.
Now Mrs. MacCarthy is pleased to have a dishwasher.

Kim fell from the ski lift.
Now both legs are in plaster.

Lisa flies to the Caribbean for her holiday.
All her colleagues are envious.

Nicole would like to return to her job after her maternity leave.
The offer of a computing course also sounds interesting.

Anne left for Paris three months ago.
Since then Thomas has a high telephone bill.

Annette is going to wallpaper her flat.
Her old pullover and working trousers are already laid out .

Carmen is giving a lecture.
Her hands are trembling.

Coherent/Incohesive

Cassie cannot concentrate on her work today.
The tray fell down twice.

Judith's heart starts pounding when she gives a talk.
A cold sweat breaks out.

Martine went to sleep on the beach.
Back and thighs are completely red.

Mrs. Ford is having visitors.
Fresh towels are put into the bathroom.

Sally is waiting to hear from Michael whom she fancies.
The heart starts beating faster every time the phone rings.

The children are playing football on the patio.
On the ground floor a window pane crashes.

Barbara leaves the lecture hall after her exam completely exhausted.
The friends are waiting at the door with a bottle of champagne.

Before the lecture the transparencies fell on the floor.
The proper order had to be assembled again.

Eric has to do a night shift at the hospital.
A strong coffee is essential.

Frank hides his purchase from his wife.
Next week is the wedding anniversary.

Global warming will cause glaciers to melt.
The sea level will rise.

In the autumn the days are getting shorter.
Elly has the bike light checked.

It is nearly 10 pm.
Marjorie turns down her stereo.

Jenny buys a bottle of champagne.
A friend is celebrating his 21st birthday.

Many heavy lorries drive past the house.
The glasses often rattle in the cupboard.

Petrol fumes are highly inflammable.
Smoking is strictly prohibited in filling stations.

The chemistry exam will take place soon.
All the textbooks are out on loan.

The cupboard is tilted.
The doors swing open by themselves.

The damage from the accident was greater than was anticipated.
The insurance company will increase the premium.

The epicentre of the earthquake was near Los Angeles.
Several houses collapsed.

The headlights are still on.
Tomorrow morning the car will probably not start.

The markets of the future lie in Eastern Europe.
Many language schools are offering Slavic languages.

The railway line at the back of the house has been shut down.
It has become much quieter in the flat.

Toothbrushes cannot easily get into the gaps between the teeth.
The use of a medicinal mouthwash is recommended.

Incoherent/Cohesive

A big anti-war demonstration has been announced.
Thus it was bitterly cold in the school.

In the city there are many offices available.
Thus not a single Christmas parcel reached its recipient on time.

John doesn't recognise who is waving to him in the park.
Luckily, his season ticket is still valid.

Last night there was a terrible storm.
This morning the office smells like a pub.

Maria is a single mother.
Therefore, water is being poured on the fire.

Marjorie has just learned to swim.
Nobody had forgotten her birthday after all.

Martin's thoughts drifted off for a moment.
His visitors are expected the following day.

Mr. Gardner has a bad cold.
His knees are grazed all the time.

Mr. Simmons takes his newspaper to work.
His colleague has given birth to a baby girl.

Mrs. Fine hopes to have a little girl.
A short walk gives her new energy.

Mrs. Jones has a hair cut.
The teacher recommends a book to her about modern parenting.

Outside it is cold and wet.
To prevent this, Martin carries digestive tablets.

Rob has booked tickets for the opera.
Now the telephone is ringing constantly.

The account was overdrawn by a lot.
As a result the village hall burned down.

The bills have not been paid for a long time.
Now all the paths have turned into mud.

The computer room has security camera surveillance.
Their frozen pizza is on special offer.

The flight has been cancelled at short notice.
Therefore Tom puts his dustbin outside.

The fountain pen has stopped working.
Therefore suntan lotion was nearly sold out.

The new glue is not very good.
Now its windows have to be de-iced.

The new government has increased the tax on energy.
That's why there were long queues at the box office.

The radio is rarely switched on.
The curtains in front of it are fluttering constantly.

The surgery involved much blood loss.
Since then the income of the local bakery has gone down.

The travel news announced long traffic jams on the motorway.
Therefore Simone has armed herself with insect repellent.

Incoherent/Incohesive

A new supermarket has been opened.
A lot of gowns had to be cleaned.

All advance tickets were sold out after a few days.
Petrol prices will go up.

In Scotland there are lots of midges.
Fred left an hour earlier.

It has been raining for days.
The phone company will disconnect the phone line.

Maggie has placed an ad in the paper.
Julia rings her baby sitter.

Marilyn has problems with her son.
Tomorrow there will be a staff outing.

Matt buys a bouquet of flowers.
The 45-minute train trip can be rather boring.

Matthew has just learned to roller skate.
In the flat handkerchiefs are lying around everywhere.

On Guy Fawkes' night a rocket exploded too early.
The automatic cash dispenser swallowed the card.

Robert heard on the weather forecast that there would be black ice on the roads.
The glasses are lying on the desk at home.

The ash trays were not emptied last night.
The roads are covered in branches.

The bloated feeling after a meal is very unpleasant.
When entering the flat Max has to take off his glasses.

The campers want to go to bed.
Every penny counts.

The car is speeding round the corner.
The tooth looks really enormous.

The car was left on the driveway for a long time.
The wallpaper came off the wall.

The electricity has failed in the shop on the corner.

Theft should be prevented.

The heating broke down over the weekend.

The whole town is swarming with police.

The postal service was on strike.

The rents are relatively low.

The summer was very hot.

The cartridge is presumably empty.

The window is not quite shut.

The volume control is no longer working.

Toby tidies up his flat.

There was nearly a collision.

Tomorrow is the day when the garbage is collected.

Toby is sitting in the departure lounge swearing to himself.

Vera cannot concentrate on her work.

Pink baby clothes are bought.

Odd-One-Out Task (used in Experiment 4)

Please circle the word that does not go with the others as quickly as you can. You have a maximum of 5 minutes, but if you finish earlier let me know IMMEDIATELY.

| | | | | |
|----|-----------|------------|----------|------------|
| 1 | blackbird | sparrow | eagle | penguin |
| 2 | plum | lettuce | radish | leek |
| 3 | calf | stallion | pony | foal |
| 4 | anchor | sail | deck | tray |
| 5 | scissors | needle | thread | excursion |
| 6 | cupboard | box | tube | drawer |
| 7 | bulb | fruit | oven | sun |
| 8 | door | window | roof | road |
| 9 | paraffin | marble | petrol | coal |
| 10 | chicken | egg | omelette | horse |
| 11 | bath | cup | suitcase | vase |
| 12 | wind | gold | oxygen | carbon |
| 13 | bush | pine cone | flower | toad |
| 14 | headlight | tail light | blood | tomato |
| 15 | cow | badger | sheep | pig |
| 16 | goal line | concert | ball | stand |
| 17 | caramel | vinegar | lemon | grapefruit |

| | | | | |
|----|-------------|--------------|-----------|-------------|
| 18 | car | flea | ball | kangaroo |
| 19 | fish | crab | cat | duck |
| 20 | psalm | chorus | song | choir |
| 21 | castle | palace | balcony | estate |
| 22 | café | pub | shop | restaurant |
| 23 | nylon | sand | plastic | popcorn |
| 24 | sandal | scarf | boot | shoe |
| 25 | foot | shoe | arm | hand |
| 26 | bathe | fly | swim | surf |
| 27 | machinegun | hand grenade | revolver | rifle |
| 28 | language | chess | golf | handball |
| 29 | couch | stretcher | hospital | mattress |
| 30 | painting | statue | drawing | picture |
| 31 | screwdriver | hammer | fence | drill |
| 32 | hedge | grass | leaf | newspaper |
| 33 | passport | rucksack | birthday | ticket |
| 34 | pedestrian | lunchbox | notebook | pencil case |
| 35 | spectator | conductor | guitarist | soprano |
| 36 | ruby | sapphire | diamond | pearl |

| | | | | |
|----|----------|---------|----------|--------------|
| 37 | brandy | cream | cider | orange |
| 38 | wheel | hammer | knife | drill |
| 39 | talk | lecture | letter | conversation |
| 40 | football | tennis | marathon | basketball |

Free Association Task (used in Experiment 4)

“When you turn over the page you will see a word. I would like you to read the word and to state as many words that are related to its meaning as you can. You will have a total of one minute to spend on each word. There are six words in total.”

Items:

Sweet

Window

Sleep

Doctor

Chair

Smell

Synonym Judgement Task (used in Experiment 4)

Instructions:

“You will be shown some phrases like this one

“slim means thin”

If you think it is true press YES

If you think it is false press NO

Remember to keep your fingers on the buttons

When you are ready to practice, press the space bar.”

Practice trials:

1. difficult/hard
2. loud/noisy
3. tall/fat
4. angry/red

Experimental Trials (randomised):

1. package/parcel
2. sick/ill
3. emotions/feelings
4. small/little
5. make/create
6. cut/slice
7. explore/investigate
8. centre/middle
9. cautious/careful
10. couple/pair
11. bald/ugly
12. careless/excited
13. lonely/stupid
14. feather/light
15. pretend/listen
16. sleep/tired
17. float/magic
18. flower/honey
19. happy/funny
20. strange/loving

Hypernym Judgement Task (used in Experiment 4)**Instructions:**

“You need to think about types of things
 You will be shown some phrases like this one

“lemonade is a type of drink”

If you think it is true press YES.

If you think it is false press NO.

Remember to keep your fingers on the buttons

When you are ready to practice, press the space bar.”

Practice trials:

1. sugar/fruit
2. cloud/cotton
3. breakfast/meal
4. London/city

Experimental Trials (random order):

1. car/vehicle
2. daisy/flower
3. fork/cutlery
4. gun/weapon
5. dictionary/book
6. diamond/jewel
7. history/subject
8. medal/award
9. fly/insect
10. hammer/tool
11. perfume/medicine
12. radio/place

13. grass/tree
14. paper/plastic
15. toothpaste/food
16. death/disease
17. beach/weather
18. cheese/meat
19. puppet/living thing
20. bread/vegetable

Same/Different task (used in Experiment 4)

Instructions:

“Are the words the same or different?

Press yes or no

When you are ready press the space bar”

Experimental trials (random order):

1. car/car
2. daisy/daisy
3. radio/radio
4. weapon/weapon
5. dictionary/dictionary
6. vehicle/vehicle
7. history/history
8. award/award
9. fly/fly
10. hammer/hammer
11. perfume/cat
12. vegetable/pull
13. beautiful/truck
14. flower/plastic
15. toothpaste/cutlery
16. death/cloud

17. bread/diamond
18. cheese/tree
19. puppet/book
20. perfume/grass