THE DOMAIN SPECIFICITY OF THE RESOURCES
REQUIRED FOR SENTENCE PROCESSING

by

MICHELLE KENNEDY

B.A., The University of Ottawa, 1998

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in
THE FACULTY OF GRADUATE STUDIES
(School of Audiology and Speech Sciences)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December, 2000

© Michelle Kennedy, 2000
In presenting this thesis in partial fulfilment of the requirements for an advanced
degree at the University of British Columbia, I agree that the Library shall make it
freely available for reference and study. I further agree that permission for extensive
copying of this thesis for scholarly purposes may be granted by the head of my
department or by his or her representatives. It is understood that copying or
publication of this thesis for financial gain shall not be allowed without my written
permission.

Department of  Audiology and Speech Sciences
The University of British Columbia
Vancouver, Canada

Date  December 15, 2000
ABSTRACT

Working memory has been proposed as the cognitive resource system that supports storage and processing while a task is being carried out. The present study investigates the nature of the resources underlying working memory tasks in different cognitive domains and how they relate to sentence processing. Specifically, it aims to determine whether the predictive ability of a working memory task relies more on similarity of cognitive domain or on similarity of the sequence of demands to the sentence processing task.

Thirty young normal participants read sentences varying in syntactic complexity in a self-paced reading experiment. Working memory was assessed using Daneman and Carpenter’s reading span task (1980), Waters and Caplan’s reading span task (1996a) and two new mathematical and visuo-spatial tasks designed to match the sequence of demands in subject-object and object-subject relative clause sentences. Correlation analyses examined the relationship between performance on each working memory task and sentence comprehension.

The results replicate previous findings of a main effect on reading time of syntactic complexity and of syntactic ambiguity for reduced relative sentences. Complexity effects were observed for subject-object conditions on the mathematical and visuo-spatial tasks. Significant correlations were observed between critical regions of subject-object sentences and (a) mathematical operations, (b) visuo-spatial operations, and (c) one version of the Waters and Caplan reading span task. This pattern of results cannot be explained by theories which incorporate a dissociation between verbal and nonverbal processing resources. The results support a connectionist approach in which familiarity with sequences of storage and processing demands across tasks explains individual differences in processing ability.
# TABLE OF CONTENTS

Abstract........................................................................................................................................... ii

Table of Contents........................................................................................................................... iii

List of Tables................................................................................................................................... ix

List of Figures................................................................................................................................. xi

Acknowledgements........................................................................................................................ xii

CHAPTER 1: LITERATURE REVIEW................................................................................................. 1

1.1 Introduction ............................................................................................................................... 1
1.2 The Study of Working Memory in Relation to Language Processing ........................................ 2
1.3 The Baddeley-Hitch Model of Working Memory ...................................................................... 3
1.4 Capacity Theory of Language Comprehension (Just & Carpenter) ........................................... 6
1.5 Separate Sentence Interpretation Resource Theory (Caplan & Waters) ................................. 10
1.6 Connectionist Views on Working Memory and Language Processing 
(e.g., MacDonald & Christiansen) .................................................................................................. 14
1.7 Summary ................................................................................................................................. 23
1.8 What do Working Memory Tests Measure? ............................................................................. 24
1.9 Summary ................................................................................................................................. 29

1.10 Research Questions and Hypotheses ....................................................................................... 30

1.10.1 Sentence Processing ............................................................................................................. 32
1.10.2 Working Memory ................................................................................................................ 33
1.10.3 Sentence Processing and Working Memory.......................................................................... 35

CHAPTER 2: METHOD....................................................................................................................... 37

2.1 Overview .................................................................................................................................. 37
2.2 Participants ............................................................................................................................... 37
2.3 Materials and Procedure ................................................................. 37
  2.3.1 General Procedure ................................................................. 37
  2.3.2 Measure of Sentence Processing ................................................. 38
  2.3.3 Sentence Complexity Variables .................................................. 40
    2.3.3.1 Introduction ................................................................. 40
    2.3.3.2 Syntactic Complexity ..................................................... 40
    2.3.3.3 Materials ................................................................. 42
    2.3.3.4 Syntactic Ambiguity ..................................................... 44
    2.3.3.5 Materials ................................................................. 45
    2.3.3.6 Fillers ................................................................. 47
    2.3.3.7 Procedure ................................................................. 47
  2.3.4 Working Memory ................................................................. 48
    2.3.4.1 Daneman and Carpenter’s Reading Span Task ....................... 50
    2.3.4.2 Waters and Caplan’s Reading Span Task .............................. 51
    2.3.4.3 Mathematical Operations ............................................. 53
    2.3.4.4 Visuo-spatial Operations ............................................. 56
  2.4 Analysis ....................................................................................... 60
    2.4.1 Syntactic Complexity .......................................................... 60
    2.4.2 Syntactic Ambiguity .......................................................... 61
    2.4.3 Daneman and Carpenter’s Reading Span Task ............................. 62
    2.4.4 Waters and Caplan’s Reading Span Task ................................. 62
    2.4.5 New Working Memory Tasks .............................................. 63
      2.4.5.1 Mathematical Operations ........................................... 63
2.4.5.2 Visuo-spatial Operations

CHAPTER 3: RESULTS

3.1 Overview

3.2 Preparation of the Data for Analysis

3.3 Syntactic Complexity and Syntactic Ambiguity

3.3.1 Syntactic Complexity: Object-Subject vs. Subject-Object Sentences

3.3.2 Syntactic Ambiguity

3.3.2.1 Main Verb Sentences

3.3.2.2 Reduced Relative Sentences

3.4 New Working Memory Tasks

3.4.1 Mathematical Operations

3.4.2 Visuo-spatial Operations

3.5 Correlations among the Working Memory Tasks

3.5.1 Critical Regions

3.5.2 Aggregate Scores

3.6 Correlations between Working Memory Measures and Specific Sentence Regions

3.6.1 Syntactic Complexity

3.6.1.1 Correlations with Critical Regions of Working Memory Measures

3.6.1.1.1 Mathematical Operations

3.6.1.1.2 Visuo-spatial Operations

3.6.1.1.3 Reading Span Tasks
3.6.1.2 Correlations with Aggregated Working Memory Measures ...85

3.6.2 Syntactic Ambiguity .................................................................86

  3.6.2.1 Correlations with Critical Regions of Working Memory Measures .........................86

  3.6.2.2 Correlations with Aggregated Working Memory Measures ................................88

3.7 Regression of the Working Memory Measures on Performance on Complex Sentences ..................89

3.8 Question Accuracy and Decision Times .....................................90

  3.8.1 Accuracy of Performance on the Comprehension Questions .........................90

  3.8.2 Response Times to the Comprehension Questions ...............................92

3.9 Summary of Results .................................................................93

CHAPTER 4: DISCUSSION .................................................................94

4.1 Introduction .................................................................................94

4.2 Overall Performance on Sentence Processing ..................................95

  4.2.1 Syntactically Complex Sentences ..............................................95

  4.2.2 Syntactically Ambiguous Sentences .........................................96

    4.2.2.1 Main Verb Sentences ......................................................96

    4.2.2.2 Reduced Relative Sentences .........................................97

4.3 Overall Performance on the New Working Memory Tasks ..................98

  4.3.1 Mathematical Operations .....................................................98

  4.3.2 Visuo-spatial Operations ......................................................99

4.4 Correlations among the Working Memory Measures ..........................100
APPENDIX G: Participants’ Scores on Standard Working Memory Measures .......... 152
APPENDIX H: Participants’ Mean Scores on Mathematical and Visuo-spatial Operations Working Memory Tasks (in msec) ........................................... 153
LIST OF TABLES

Table 1. Syntactic Complexity Variable: Group (N = 30) Response Time Means and
Standard Deviations by Region (in msec) ..........................................................67

Table 2. Syntactic Ambiguity Variable: Group Response Time Means and Standard
Deviations by Sentence Regions (in msec) ......................................................70

Table 3. New Working Memory Tasks: Group Response Time Means and Standard
Deviations by Region for the Mathematical Operations and Visuo-spatial
Operations Tasks (in msec) ........................................................................75

Table 4. Correlation Matrix of Performance on Working Memory Measures ..........81

Table 5. Correlations of Performance on all Working Memory Measures Using Aggregate
Scores for Mathematical Operations and Visuo-spatial Operations .................82

Table 6. Correlations between Working Memory Measures (Critical Regions of
Mathematical and Visuo-spatial Operations) and Critical Regions of the
Object-subject and Subject-object Sentences ...............................................83

Table 7. Correlations between Mathematical and Visuo-spatial Operations Aggregate Scores
and Critical Regions of the Object-subject and Subject-object Sentences .........86

Table 8. Correlations between the Working Memory Measures and Regions 2 and 3 of the
Main Verb and Reduced Relative Ambiguous and Unambiguous Sentences....87

Table 9. Correlations between Mathematical and Visuo-spatial Aggregate Scores and
Regions 2 and 3 of the Main Verb and Reduced Relative Ambiguous and
Unambiguous Sentences ..............................................................................88
Table 10. Summary of Stepwise Regression Analysis for Variables Predicting Ability to Process Complex Subject-object Sentences ......................................................... 90

Table 11. Overall Accuracy on Sentence Comprehension Questions and Working Memory Tasks.............................................................................................................. 91
LIST OF FIGURES

Figure 1. Reaction Times by Region for Object-subject and Subject-object Sentences ................................................................. 68

Figure 2. Subject-object Minus Object-subject ......................................................................................................................... 69

Figure 3. Reaction Times by Region for Main Verb Ambiguous and Unambiguous Sentences ......................................................... 71

Figure 4. Main Verb Ambiguous Minus Unambiguous ............................................................................................................. 72

Figure 5. Reaction Times by Region for Reduced Relative Ambiguous and Unambiguous Sentences .................................................. 73

Figure 6. Reduced Relative Ambiguous Minus Unambiguous .................................................................................................. 74

Figure 7. Reaction Times by Region for Mathematical Operations (OS and SO conditions) ............................................................ 76

Figure 8. Mathematical Operations: Subject-object Minus Object-subject ................................................................. 76

Figure 9. Reaction Times by Region for Visuo-spatial Operations (OS and SO conditions) ............................................................. 78

Figure 10. Visuo-spatial Operations: Subject-object Minus Object-subject ................................................................. 78
ACKNOWLEDGEMENTS

The completion of this work would not have been possible without the help and support of the faculty and staff of the School of Audiology and Speech Sciences, my classmates, friends and family.

Special thanks are due to Dr. Jeff Small, whose patience, guidance and sense of humour were critical to keeping this endeavour on track. In particular, I thank him for reminding me to 'enjoy the process' of thesis writing. Thanks also to Dr. Kathy Pichora-Fuller, for her insight and advice.

Many thanks to my 'thesis friends', Martha and Carole, for the hours spent encouraging one another and commiserating, and also to Rob, Beth and Gigi, for their unwavering support and diligence in making sure I took time off to relax.

And finally, thanks to my mom and dad, and sister Suzanne, who have supported me in so many ways throughout my six years of university. None of this would have been possible without them.
CHAPTER 1: LITERATURE REVIEW

1.1 Introduction

Over the last twenty years, considerable research has focused on the nature of the working memory system. Today, this remains a controversial area with respect to the underlying nature of working memory, how it should be measured and how it relates to language. An early and influential model of working memory (Baddeley & Hitch, 1974) posits a central executive that allocates attention and processing resources, and two slave systems of the central executive, the articulatory loop and the visuo-spatial sketchpad. Other researchers (e.g. Just & Carpenter, 1992) find lack of motivation for a separate central executive and hold that there is a unitary set of resources that are drawn upon for processing and storage for language. Others (e.g. Caplan & Waters, 1999) propose still greater fractionation of the working memory system for language into sentence-specific processing, as opposed to other verbally mediated tasks. A more recent proposal (MacDonald & Christiansen, in press) is that individual differences on various working memory tests reflect varying experience with language. The current study proposes to further examine the nature of the resources underlying working memory tasks in various cognitive domains, and how they relate to sentence processing as it unfolds over time. Specifically, it will investigate whether the resources supporting sentence processing are domain specific.

There are various theoretical approaches to conceptualizing working memory for language, which have been touched upon above. These approaches vary according to the focus of the researchers who developed them. For example, Baddeley and Hitch developed a model of working memory based on neuropsychology (1974). Just and Carpenter’s view of working memory focuses on individual differences, which they have simulated using
symbolic computational models (e.g., Just & Carpenter, 1992; Haarman, Just & Carpenter, 1997). Caplan and Waters’ conceptualization of working memory has been shaped by their modular view of language processing (1999). MacDonald and Christiansen take a distributed connectionist view of language processing, and thus of working memory. These different approaches to the concept of working memory as it supports language processing will be reviewed in the following sections. Tests of working memory, which have been developed based on these models, will also be discussed. This review is limited to those models of working memory that make claims specific to sentence processing (for an overview of working memory models relating to cognition in general, see Miyake and Shah, 1999).

1.2 The Study of Working Memory in Relation to Language Processing

“Working memory” is a relatively recent inception in the scientific world. Just and Carpenter provide a brief historical account of working memory (1992). It had its beginnings in short-term memory, which was conceived of to account for the storage of information over short periods of time until it needs to be recalled. A classic example of short-term memory use cited by Just and Carpenter (1992) is the storage of a telephone number for the brief interval between looking it up in the phone book and dialing the phone. Short-term memory was later expanded to include functions of keeping information activated for elaboration in long-term memory, and for keeping partial results active in memory during processing. The term 'working' was added to memory to accommodate the computational functions it took on in the first model of working memory (Baddeley & Hitch, 1974).
1.3 The Baddeley-Hitch Model of Working Memory

In the early seventies, Baddeley and Hitch began to develop what would become the first model of working memory. Their idea for the model was based on a series of studies on the memory used for reasoning, language and learning (Baddeley & Hitch, 1974). The researchers used a dual task paradigm, in which subjects were asked to maintain a series of digits in memory (the typical short term memory test) while performing tasks involving either reasoning, language comprehension or learning. They found that performance across tasks deteriorated as the number of digits to be maintained in memory increased. However, performance did not deteriorate entirely, even when the number of digits to be recalled was at a subject’s span level. This led them to conclude that traditional short-term memory and the system they were investigating (working memory) were not the same system (Baddeley, 1996). The results of these studies led the authors to propose that “the core of the working memory system consists of a limited capacity ‘work space’ which can be divided between storage and control processing demands” (Baddeley & Hitch, 1974, p. 76).

Baddeley “favours a view of working memory as a general purpose system that can be fractionated into subsystems” (1996, p.7). Accordingly, the multiple-component model of working memory developed by Baddeley and Hitch (1974) is a general system which includes three major components: the central executive, the articulatory loop (sometimes referred to as the phonological loop) and the visuo-spatial sketch pad.

The central executive is the core of the working memory system, which allocates attention and resources to processing. The articulatory loop and the visuo-spatial sketch pad are slave systems to the central executive. Only in recent years has attention turned towards investigating and defining the role of the central executive, which was previously referred to
as a “conceptual ragbag” (Baddeley, 1986, p. 24). Whereas it was first proposed that the central executive could allocate extra resources for storage when the slave systems exceeded their storage capacity, Baddeley no longer supports this view (Baddeley & Logie, 1999). The slave systems instead access long-term memory when they have reached their limit. The role of the central executive is currently thought to include control and coordination of the slave systems and the capacity to focus, switch and divide attention among tasks (Baddeley & Logie, 1999; Baddeley, 2000).

A greater area of focus in the literature has been on the two slave systems. The articulatory loop maintains phonemic information active in memory so that verbal computations can be carried out. It consists of two components: the verbal store and the articulatory rehearsal process. The verbal store holds “speech-based information represented in traces” (Baddeley, 1996, p. 17), while the articulatory rehearsal process is used to enter information or to maintain traces in the verbal store. One piece of evidence supporting the existence of a verbal store is the ‘phonological similarity effect’, the observation that it is more difficult to remember sequences of words or sounds which are phonologically similar than sequences that are not phonologically similar. Evidence supporting the articulatory rehearsal process is the ‘irrelevant speech effect’, the fact that irrelevant speech is disruptive to attempts to remember other information, such as a string of digits. The other slave system, the visuo-spatial sketch pad, specializes in the storage and processing of visual and spatial imagery. Recent dual task studies have found that visual, but not spatial, stimuli disrupt measures of visual imagery, whereas spatial tracking (e.g., pointing to the source of a sound), but not visual tasks, disrupts spatial memory tasks (Baddeley, 1996). Based on such evidence, it has been suggested that this system can be further fractionated into two
dissociated subsystems, the visual cache and the inner scribe, which support working
memory for visual patterns and for sequences of movements, respectively (Baddeley &
Logie, 1999). Both slave systems have limited resources at their disposal; when they reach
their limit, further resources must be drawn from long-term memory or other subsystems to
support storage or processing. Working memory is considered to underlie any task involving
"on-line cognition – the moment-to-moment monitoring, processing and maintenance of
information" (Baddeley & Logie, 1999, p. 28). This includes, for example, performing
certain mental mathematical calculations, reading comprehension, and reasoning.

The Baddeley-Hitch model of working memory has been of particular interest for
understanding language comprehension processes. Although researchers long suspected that
short term memory (i.e., the articulatory loop) was an important factor underlying reading
comprehension, traditional measures of short term memory, such as word span or digit span,
were paradoxically not predictive of reading ability because these measures do not account
for processing. Therefore, Daneman and Carpenter (1980) developed a test of working
memory for language based on the Baddeley-Hitch model, which included a processing
component as well as a storage component. The test required subjects to read sets of
sentences aloud and at the end of each set, to recall the sentence-final words in order. Sets
increased in size (from two to six sentences) until the subject could no longer recall all the
words in a set. The highest level at which a subject recalled all words correctly on two out of
three sets was defined as the subject’s reading span. Reading span was found to correlate
highly with measures of reading ability, such as the Verbal Scholastic Aptitude Test (VSAT),
with a correlation of .59, and specific language processing measures such as pronominal
reference, with a correlation of .90 (Daneman & Carpenter, 1980). Numerous correlational
studies have since been published which further establish the relationship between various versions of the original Daneman-Carpenter reading span test and a wide range of language processing tasks, from standardized tests of language ability (e.g., the General Aptitude Test Battery: Vocabulary, the Nelson-Denny Reading Test) to more specific measures (e.g., making inferences, detecting ambiguity) (for a review see Daneman & Merikle, 1996). The proliferation of experimental studies in this area stimulated thought regarding the nature of the working memory system and its relationship to language processing, giving rise to new models of working memory.

1.4 Capacity Theory of Language Comprehension (Just & Carpenter)

In a 1992 paper, Just and Carpenter advanced a model of working memory for language comprehension. These researchers proposed that individuals differ in terms of capacity, and that capacity constrains comprehension. They advanced a unitary view of working memory for language comprehension, corresponding roughly to the part of Baddeley's central executive that is concerned with language comprehension. Just and Carpenter's model of working memory is based on activation, which they describe as the "resource" underlying storage and processing. According to this model, individuals have different capacities, or total amounts of activation. There is an activation level associated with each representational element (such as a word, grammatical structure, thematic role, etc.). Elements become activated during the course of language comprehension. When a certain threshold of activation level is reached, that representational element comes into working memory from long-term memory, and is available for processing. However, if the number of elements in working memory exceeds the total activation available to the system, then elements activated earlier can be bumped out of working memory, in a sort of "forgetting by displacement"
process (Just & Carpenter, 1992, p. 123). In this way, earlier partial representations may be forgotten, and are thus not available for integration with later sentential elements. The authors describe a “trading relation” between storage and processing, such that when demand for activation exceeds supply, activation for both storage and processing is cut back proportionately to the amount being used by each component (Just & Carpenter, 1992, p.123). The result is that demanding tasks are slower and more vulnerable to decay of partial results. With respect to language comprehension, this translates into differences in speed and accuracy of comprehension as a function of an individual’s capacity, or total activation available.

Much research lends support to Just and Carpenter’s capacity constrained model of language comprehension (see Daneman & Merikle, 1996 for a review). For example, a study by King and Just (1991) separated subjects into high and low working memory span, using the Daneman and Carpenter (1980) reading span test, and used an on-line measure to explore subjects’ processing of object-relative and subject-relative sentences (e.g., The reporter that the senator attacked admitted the error and The reporter that attacked the senator admitted the error, respectively). The authors found a significant interaction between working memory span and sentence type in line with the predictions of the capacity theory. Namely, high span subjects were faster and more accurate than low span subjects in processing the more demanding object-relative sentences.

Another study by MacDonald, Just and Carpenter (1992) tested the predictions of the model using comprehension of syntactically unambiguous versus ambiguous sentences. The ambiguous sentences were constructed to take either a main verb interpretation (e.g., The experienced soldiers warned about the dangers before the midnight raid), or a reduced
relative interpretation (e.g., *The experienced soldiers warned about the dangers conducted the midnight raid*). Using a self-paced word-by-word presentation, the researchers found a significant interaction between span and ambiguity; however, it was the high span subjects who took significantly longer to process the ambiguous than unambiguous sentences. This seemingly counterintuitive finding could be explained by the capacity constrained model by proposing that the high span subjects’ longer response times were due to maintaining multiple interpretations of the sentence during processing. The result was greater accuracy in the processing of ambiguous sentences that resolved with a main verb interpretation at the cost of slower processing. Low span readers, who lacked sufficient activation to maintain multiple representations, were faster at processing but much less accurate on answering comprehension questions. Sentences with a relative clause interpretation (e.g., *The experienced soldiers warned about the dangers conducted the midnight raid,* versus *The experienced soldiers who were told about the dangers conducted the midnight raid*) revealed a similar pattern of results as the sentences with main verb resolution; there was a significant interaction between span and ambiguity, with high span subjects taking significantly longer to process the ambiguous than the unambiguous sentences. Sentences with reduced relative clauses appeared to be more demanding than those with main verb resolution; both high and low span subjects had longer reading times and lower accuracy on comprehension questions for reduced relative sentences as compared with sentences with main verb resolution. It should be noted that these findings have since been reinterpreted in a different light by the first author and other researchers (Pearlmutter & MacDonald, 1995 — see section 1.6).

Further support for Just and Carpenter’s model of working memory comes from a simulation model they constructed called CC Reader (Capacity Constrained Reader) (1992).
This model is a hybrid of a symbolic production system and a connectionist system. It was constructed following the parameters set out by the capacity constrained model: there is a total amount of activation, activation thresholds are specific to representational units, and an intermediate trading relation exists between storage and processing (such that both are reduced proportionately if total activation is exceeded). By varying the amount of total activation in the model, they were able to simulate high and low span subjects’ performance on object-relative and subject-relative sentences as reported for humans in King and Just’s (1991) study. This finding has more recently been replicated in Haarman, Just and Carpenter (1997), using an improved version of CC Reader to mimic aphasic sentence comprehension.

Other evidence comes from Just, Carpenter and Keller (1996), who conducted a study involving brain imaging (fMRI). The results of this study indicated that overlapping areas (particularly Wernicke’s area) were activated during both reading alone and reading-and-maintaining (similar to the reading span test) conditions. Significantly higher overall activation in these areas during the read-and-maintain condition is interpreted by these authors as support for the theory that both processing and storage are central to working memory, and that the reading span test taxes this system.

These studies are a sample from many that have provided evidence in support of Just and Carpenter’s model of working memory as it relates to sentence processing. While much evidence exists in support of the capacity theory, other researchers have failed to replicate Just and Carpenter’s (and others’) results (Waters & Caplan, 1996b; Kemtes & Kemper, 1997; Caplan & Waters, 1999).
1.5 Separate Sentence Interpretation Resource Theory (Caplan & Waters)

Caplan and Waters have entered into a long-running debate with Just and Carpenter regarding the nature of the resources underlying sentence processing. Much of their critique of Just and Carpenter is based upon a failure to replicate previous research and the apparent incompatibility of the capacity theory with neuropsychological data. From these criticisms, Caplan and Waters developed their own theory of sentence processing capacity, to be described below.

Caplan and Waters (1999) take issue with the claim of Just and Carpenter that a single pool of resources underlies both sentence processing and other verbally mediated tasks. Caplan and Waters subscribe to a modular view of language comprehension; thus, syntactic parsing is seen as separate and encapsulated from other verbal processes. Following from this view of language, Caplan and Waters divide verbal working memory into two subsystems which serve interpretive and post-interpretive sentence processing. According to these researchers, the interpretive working memory system is responsible for parsing the acoustic signal and assigning meaning, which includes the processes involved in sentence comprehension. Post-interpretive processes, on the other hand, make use of the interpretive meaning in other verbally mediated tasks, such as reasoning and planning for action (e.g., sentence-picture matching). Based on this distinction, Caplan and Waters posit the separate sentence interpretation resource (SSIR) theory. Because interpretive and post-interpretive operations are separate, the efficiency with which operations are carried out under each may not necessarily be correlated. A central prediction of the SSIR, then, is that performance on general verbal working memory tasks will not correlate with the efficiency of interpretive sentence processing. Thus, the SSIR predicts that no interactions will be found between
Daneman and Carpenter’s (1980) reading span and sentence complexity. Indeed, one of the criticisms that Caplan and Waters have levelled at Just and Carpenter’s theory is their inability to replicate several of the results that support the capacity theory. Caplan and Waters (1999) report that they have been unable to replicate King and Just’s (1991) finding of an interaction between span, syntactic complexity and region of the sentence using an online auditory moving window technique. They have also failed to replicate the span × ambiguity interaction found by MacDonald et al. (1992). Another prediction of the SSIR is that there should be no interaction between external memory load and sentence complexity since interpretive processing is thought to be separate from processing external to this subsystem. External memory load can be imposed in a dual task where sentence processing is carried out under a concurrent task of recalling a set of digits or words (usually at the subject’s span, or span plus one) or by incorporating an increasing memory load into a sentence processing task, like the reading span task (e.g., King & Just’s 1991). Caplan and Waters (1999) report on previous studies they have conducted using concurrent memory load with sentence-picture matching, enactment of thematic roles and plausibility judgements (all of which are off-line measures of sentence processing). In these studies, the authors failed to find an interaction between external memory load and sentence complexity, such that increases in external memory load did not negatively impact more on sentence processing for complex sentences than on simple sentences. Given that Caplan and Waters’ findings support the null hypothesis, they provide, at best, weak support for the SSIR theory.

More direct support of the separate sentence interpretation resource comes from neuropsychological studies. Waters, Caplan and Hildebrandt (1991) provide a case study of a patient, B.O., with short-term memory deficits in articulatory rehearsal and possibly the
phonological store (referring to the Baddeley-Hitch slave system). They note that some researchers would expect that impairment in these areas would lead to impaired syntactic processing. However, B.O. was determined to have normal sentence comprehension, even for complex garden path sentences, in spite of her impaired short-term memory. This is taken to support the theory that sentence processing and other verbal tasks draw on separate pools of resources; performance on short-term memory tasks was affected by the impairment, while sentence processing was not. However, upon closer inspection of B.O.'s data, it becomes apparent that B.O. did indeed have significantly greater difficulty in comprehending object-relativized sentences, in particular, subject-object constructions. This led Waters et al. to conclude that "this might be because STM [short term memory] is involved in the comprehension of sentences with object relativisation" (1991, p.119). Thus, the authors themselves admit that processing of syntactically complex sentences was affected by short-term memory impairment, a finding which does not support the SSIR theory.

Caplan and Waters (1999) also cite evidence from patients with central executive limitations, such as those with dementia of the Alzheimer's type (DAT) in support of the SSIR theory. Although people with DAT have been reported to achieve very low scores on working memory tasks, Caplan and Waters have found that they do not perform disproportionately worse on syntactically complex sentences when measured on sentence-picture matching tasks or acceptability judgement tasks. However, data from other researchers (e.g., Small, Andersen & Kempler, 1997; Small, Kemper & Lyons, 2000) have demonstrated strong correlations between working memory capacity and comprehension of syntactically complex sentences for patients with central executive impairments.
Further evidence comes from patients with aphasia, who are considered to have reduced resources overall. Caplan and Waters report that these patients do show effects of syntactic complexity, but that this effect is not increased significantly by a concurrent digit span load. The authors take this as further support for their claim that the resources for sentence processing are not the same as those involved in maintaining an external memory load. Data from various patient populations provides some positive evidence in support of the separate sentence interpretation resource theory, although the counter-evidence cannot be readily dismissed.

Christiansen and MacDonald (1999a) offer some criticisms of Caplan and Waters’ proposed separate sentence interpretation resource theory. First, they suggest that under the Caplan and Waters’ view, the number of working memory systems could multiply indefinitely. Caplan and Waters suggest that since on-line sentence comprehension and performance on post-interpretive tasks (e.g., following verbal instructions) do not correlate, then separate pools of resources underlie the two. By this logic, Christiansen and MacDonald claim that any two tasks that do not correlate will necessitate separate working memories and that the “decision to have one vs. two vs. twenty working memories is unprincipled” (1999a, p. 98). Furthermore, Christiansen and MacDonald point out that many of the studies reported by Caplan and Waters (1999) involve off-line measures of sentence processing (e.g., picture matching), and that neuropsychological data may offer false support for the SSIR, only because so few on-line studies with patient populations exist.

An earlier criticism of resource theories in general, which applies to both the capacity theory and the separate sentence interpretation resource theory, comes from Navon (1984). Navon is sceptical of “resource theory”, which makes claims based on pools of resources
underlying processing. According to Navon, resources are a "theoretical soup stone", that is, individual differences could be accounted for equally well without recourse to resources. For example, resource theory makes the claim that two tasks carried out concurrently will interfere with one another if they are drawing on the same resource. However, as Navon points out, a task may create outputs or throughputs or other side-effects that interfere with performance on another task, regardless of whether they are drawing on the same resource. Navon also underlines the fact that the notion of resources is unfalsifiable; when the theory's predictions are not met there are "built-in escapes", such as data limits, multiple different resource pools, processing below full capacity, etc. (Navon, 1984, p.231). These criticisms represent a serious problem for Just and Carpenter's and Caplan and Waters' models of working memory.

1.6 Connectionist Views on Working Memory and Language Processing
(e.g., MacDonald & Christiansen)

Computational connectionism takes a different approach to working memory, suggesting that constraints on cognitive processing ability result from the "architecture" of the processor itself as well as experience with language rather than a separate "working memory" capacity. According to some connectionist modellers, language processing is constraint-based and governed by probabilistic information gathered through experience with language. That is, frequency of lexical items, frequency of verb argument structures, plausibility and context guide language comprehension processes.

Connectionist models operate through the spreading of activation across connected units which are organized into nodes and layers. Weights on the connections control the spread of activation. In a somewhat oversimplified analogy, the units in these models correspond to neurons, the weights to synapses, activation to the electrical output of neurons and
connections to dendrites and axons. Parallel distributed-processing (PDP) is one approach to connectionist modelling. In contrast to many earlier computational models of sentence processing, PDP connectionist models are neither modular, serial, localist or deterministic. Content from one level of information (e.g., lexical) influences processing at another level (e.g., syntactic), and therefore the system is nonmodular. In addition, all levels of linguistic representation are processed in parallel, in contrast to serial models in which information from one level feeds into the next level in stages. The unit (or representational) structure in PDP models is distributed (as opposed to localist) such that a linguistic feature or structure is represented by patterns of activation across units, rather than by a single unit. Thus, linguistic features are organized into nodes, and processing consists of recognizing patterns of activation across nodes. In terms of architecture, PDP connectionist models are multi-layered, typically consisting of an input layer, output layer and 'hidden unit' layer(s).

Algorithms constrain how the models operate, determining such things as minimum and maximum activation levels of units, the level of activation necessary for elements to reach threshold, and decay rate of activation. Importantly, PDP models incorporate a back propagation algorithm. Back propagation allows the model to learn by comparing the actual output with the desired output in the training phase. Based on the discrepancy between the two, it modifies weights in the hidden and output layers in order to achieve a closer match between actual and desired output. Processing in PDP models is constraint-based, meaning that statistical and probabilistic factors and not absolute (deterministic) rules constrain the final output of comprehension or production. According to Seidenberg and MacDonald, "the essential idea of the constraint-based approach is that comprehending or producing an utterance involves interactions among a large number of probabilistic constraints over
different types of linguistic and non-linguistic information" (1999, p. 576). Thus, the amount of experience with language will determine sensitivity to the frequency of certain patterns of activation which in turn will determine processing ability.

Christiansen and Chater (1999) provide an example of a connectionist model of syntactic processing. These researchers implemented a simple recurrent network (SRN) to model human language processing limits for reading comprehension of recursive sentences, such as centre-embedded structures. Multiply embedded structures are next to impossible to understand, as in the following example:

*The student that the teacher that the principal commended helped studied.*

Such constructions are neither produced nor easily understood; however, there is no inherent linguistic reason why the number of recursions in a sentence should be restricted.

Christiansen and Chater (1999) point out that symbolic models have been unable to replicate patterns of human performance on such structures without imposing arbitrary processing limitations on the network, prohibiting double or multiple embeddings. Importantly, in simple recurrent networks "the current set of hidden unit values is ‘copied back’ to a set of additional input units, and paired with the next input to the network" (Christiansen & Chater, p. 164). This gives the network capability to learn from prior experience, effectively giving it a "memory for past inputs" (p. 164). Thus, the ability to learn arises from the hidden unit layer, and the ability of the network to process certain structures is related to previous experience with similar patterns of activation. Unlike previous models, Christiansen and Chater’s SRN successfully replicated human capacity constraints in processing recursive structures without setting arbitrary limits on the network. Furthermore, the SRN’s performance was not affected by changes to network parameters (e.g., the size of the hidden
unit layer) or the training corpus. They therefore conclude that their results are due to
"intrinsic constraints of the SRN architecture" (Christiansen & Chater, 1999, p. 165). The
authors note that such processing limitations are not necessarily restricted to language. As
cited by Christiansen and Chater, an earlier study by Larkin and Burns (1977) revealed that
non-linguistic material yielded similar patterns of degradation with increasing levels of
embedding. Larkin and Burns presented participants with sequences of digits, or letters and
digits, or nouns and verbs following the pattern in recursive sentence structures. For
example:

<table>
<thead>
<tr>
<th>The student</th>
<th>that</th>
<th>the teacher</th>
<th>that</th>
<th>the principal</th>
<th>observed</th>
<th>helped</th>
<th>studied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>Noun</td>
<td>Noun</td>
<td>Verb</td>
<td>Verb</td>
<td>Verb</td>
<td>Digit</td>
<td>Digit</td>
</tr>
<tr>
<td>Letter</td>
<td>Letter</td>
<td>Letter</td>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
</tr>
<tr>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
<td>Digit</td>
</tr>
</tbody>
</table>

Participants listened to sequences such as FJN371 and were required to repeat them back in
their component pairs: N3, J7, F1. The number of pairs increased from two to five.
Participants performed slightly better overall in the digit-only and digit-letter conditions than
in the noun-verb and full sentence conditions. However, the general pattern of performance
degradation was similar across conditions as the number of embeddings increased. These
findings and their own lead Christiansen and Chater to conclude that “constraints on complex
recursive structures, such as center-embedding may derive from non-linguistic processing
constraints” (1999, p. 172). Christiansen and Chater's work suggests that certain limitations
in human processing of linguistic and non-linguistic structures may be due to the structure of
the processing system itself and the nature of processing demands, rather than a separate
working memory system.
The connectionist approach to "resources" is unlike that of both Just and Carpenter and Waters and Caplan in that it does not require a separate "capacity" for processing and storage. In connectionist neural networks, there is no distinction between linguistic representations, linguistic processing and working memory capacity. Capacity, in connectionist terms, is defined simply as the "efficiency of passing activation through a network" (Christiansen & MacDonald, 1999a, p. 97). According to MacDonald and Christiansen (in press), individual differences in language processing are not based on working memory capacity, but on individual differences in linguistic experience (and differences in architecture, e.g., in the accuracy of phonological representations). These authors argue that there is an "artificial" distinction between tasks measuring language processing and those measuring verbal working memory: both measure "language processing skill". Thus, working memory measures correlate with language processing tasks not because measures of working memory tap resource capacity, but because the same architecture, processing constraints and experience underlies performance on both tasks.

MacDonald and Christiansen (in press) offer reinterpretations of several earlier findings, from a connectionist viewpoint. The differences King and Just (1991) observed between subject- and object-relatives are an example of a frequency × regularity interaction. Subject-relatives are more 'regular' with respect to word order than are object-relatives, and occur much more frequently in the language. Higher exposure to subject-relatives by low and high span subjects makes this an easier construction to understand, for both high and low span subjects. High span subjects have more experience with language overall, and thus have more exposure to the less frequent object-relative sentences. For this reason, high span subjects perform better on these 'irregular' sentences than low span subjects because they are
higher in frequency for the high span subjects, thus leading to a frequency x regularity interaction. MacDonald and Christiansen (in press) replicated King and Just’s (1991) findings in an experiment using ten simple recurrent networks. Each network began with different starting weights, reflecting the individual differences subjects bring to a language processing experiment. The networks were created with 31 input and output units, and 60 hidden and context units. Activation of an input pattern on the input level was propagated through the hidden layer to the output layer. During the training phase, the pattern of activation on the output layer was compared with the correct pattern through back-propagation. The initial weights on units in the hidden layer could then be adjusted accordingly. The context layer played a role in subsequent time steps, maintaining the activation pattern from the hidden layer in the previous time step. Training consisted of passes through a corpus of ten thousand sentences, in which subject and object relative sentences were represented in equal probability (2.5% of the corpus). These relative clause sentences were randomly presented with simple transitive and simple intransitive sentences of varying lengths. One pass through the corpus was termed one epoch. Sentence processing ability was measured on a set of test sentences after each of three epochs of training to determine whether performance after different epochs (i.e., amounts of experience) mirrored the performance of low and high span subjects, who are assumed to have different levels of experience with language. Performance of the networks was measured by having the SRN predict the probability of possible next words given an input string (e.g., the network might be given the sequence The child, and would then predict that possible next words would include a past tense verb, a third person singular present tense verb or the relative pronoun who). The measure of performance of the networks was called the “Grammatical Prediction
Error" (GPE) which takes "hits (H), false alarms (F), misses (M) as well as correct rejections into account" (MacDonald & Christiansen, in press, p. 30). GPE values can be compared to reading times, with low values corresponding to short response times, and high values corresponding to long response times. The networks performed in line with King and Just's (1991) predictions; they performed like low span subjects (with higher GPE values) early on in training, and more like high span subjects (with lower GPE values) following three epochs of training. The researchers point out that results of this simulation paralleled the span × sentence type interaction found in the CC Reader simulation (Just & Carpenter, 1992; Haarman, Just & Carpenter, 1997); GPE values revealed that training resulted in little change for more regular subject-relative sentences, but had a much larger effect for irregular object-relative sentences. Thus, individual differences in the size of a separate "resource" do not need to be called upon to explain King and Just's (1991) findings.

Further evidence supporting a connectionist view on "working memory capacity" comes from behavioural research. In another test of King and Just's (1991) findings, Christiansen and MacDonald (1999b) report that the ability of subjects to process object-relative sentences could be manipulated by increasing their exposure to this sentence type (i.e., through training). Subjects who were given concentrated experience with centre-embedded relative clause constructions showed improved performance on those sentences compared to a group of control subjects, mirroring the pattern of results that King and Just had attributed to differences in working memory capacity. On the other hand, no post-training changes were found for performance on working memory tasks.

These research findings demonstrate the role of experience in determining ability to process specific complex sentence types. In a constraint-based model of sentence processing,
experience modulates not only the ability to process specific types of sentences, but is also important with respect to the ability to use probabilistic information during on-line processing. Recent theories of syntactic ambiguity resolution have proposed that sensitivity to probabilistic information in lexical entries as well as higher-level noun-verb combinations and context are responsible for the interpretation of ambiguous sentences.

Pearlmutter and MacDonald (1995) offer a reinterpretation from a connectionist point of view of the MacDonald et al. (1992) study focusing on syntactic ambiguity (see section 1.4 for the details of this study). Pearlmutter and MacDonald created a new stimuli set, comparing main verb unambiguous (e.g., *The soup bubbled in the pot but was not ready to eat*) and ambiguous sentences (e.g., *The soup cooked in the pot but was not ready to eat*). These researchers replicated MacDonald et al.’s (1992) results; high span subjects had longer reading times at disambiguating points in ambiguous sentences than in unambiguous sentences, whereas low span subjects exhibited no differences in reading times between the two sentence types. However, rather than explain this difference in terms of ability of high span subjects to maintain multiple representations over longer distances, Pearlmutter and MacDonald interpret the difference in terms of probabilistic constraints. In a preliminary experiment, participants rated the plausibility of sentences used in the MacDonald et al. (1992) study. The main verb unambiguous sentences were found to be less plausible than the main verb ambiguous sentences. In a second experiment, participants rated the plausibility of the noun-verb combinations used in the stimuli in the 1995 study. Participants were asked to rate the plausibility of main verb transitive (e.g., *The soup had cooked the...*) versus intransitive (e.g., *The soup had cooked in the...*) interpretations, as well as relative clause (*The soup that was cooked in the...*) and unambiguous main verb (*The soup bubbled in the...*)
interpretations. A regression analysis determined that high span participants’ reading times at critical points of ambiguity resolution were predicted by plausibility ratings of alternative interpretations. This was not the case for low span participants. The authors conclude that both high and low span subjects have knowledge of contextual constraints that affect the plausibility of alternative interpretations; however, while the high span subjects are able to make use of this knowledge on-line, low span subjects are not. This constraint-based interpretation is thus able to account for the findings in the 1992 experiment without appealing to “working memory” capacity.

In response to Christiansen and MacDonald’s (1999a) re-interpretation of Caplan and Waters (1999) findings, the latter authors do not consider the connectionist view on working memory to be a step beyond resource theory. They note that elements of connectionist models are analogous to the terms they use to describe their model based on sentence processing resources. For example, certain elements determine the efficiency of a connectionist network, be it the number of hidden units, the number of connections or the strength of the connections. Connectionist models also have measures of complexity, such as the number of times activation is passed through before the response can be selected. Processing load can also be measured in connectionist networks (Caplan and Waters do not explain how this is achieved). Because of these similarities, Caplan and Waters argue that the “concept of processing resource or working memory is not eliminated in connectionist models so much as transformed because of the nature of the models” (1999, p. 115). By this view, connectionist views on working memory are not so different from previous models.

Caplan and Waters seem to overlook the fact that connectionist models computationally operationalize individual differences such as those seen in the ability to process complex
sentences. Connectionist models of human sentence processing are able to quantify and therefore better explain individual differences without recourse to vague and ill-defined “resource pools”. They are able to determine what aspects of a model influence processing ability, be it the size of the hidden unit layer or strength of connections (which comes as a result of familiarity with patterns in the input). This represents a distinct advantage over other approaches to studying individual differences, in that it allows us to operationally define “resources” for sentence processing.

1.7 Summary

The debate over the structure of working memory, or whether it even exists as a separate system, is far from over. Baddeley and Hitch’s (1974) original model of working memory, with a central executive responsible for controlling and coordinating short-term memory slave systems, has been modified by a number of researchers focusing on language processing. Just and Carpenter (1992) developed a model of working memory for language which they claim approximates that part of Baddeley’s model that is responsible for language. Caplan and Waters (1999) proposed a further fractionation of working memory resources into those underlying syntactic processing and those underlying other verbal processing. A more recent proposal has come from connectionist researchers, such as Chater and Christiansen (1999) and MacDonald and Christiansen (in press), who claim that there is no need for a separate working memory capacity, and that individual differences in linguistic and nonlinguistic processing can be explained by experience and characteristics of the processing system itself. The evidence cited on all sides of this debate, from studies of individual differences to neuropsychological data to symbolic and connectionist models suggests a need for further research. The present study is an effort to investigate the nature of
the "resources" underlying sentence processing and the extent to which processing in different cognitive domains makes use of a similar set of constraints.

1.8 What do Working Memory Tests Measure?

A second controversy has surrounded the ways in which working memory can be measured and whether working memory systems are domain-specific or general (e.g., language only or more than one domain of cognition). Considerable research has focused on the nature of processing required on working memory measures in different cognitive domains, how working memory tasks correlate with one another, and to what extent they predict performance on other cognitive tasks (e.g., language processing).

A wide range of working memory tests for different cognitive domains has been developed over the years. As Jurden (1995) points out, there is as yet no standard set of working memory measures. Indeed, there is a prevailing lack of agreement as to exactly how working memory should be measured. Most working memory tests involve both a processing and a memory component, in keeping with the Baddeley-Hitch model of working memory (1974). However, tests differ in terms of whether one component (usually memory) or both form the dependent measure; the Daneman and Carpenter (1980) reading span task is an example of the former, whereas the Waters and Caplan (1996a) reading span task is an example of the latter. The trade-off between storage and processing, and which component of working memory contributes to the predictive ability of these tests has been investigated by Turner and Engle (1989). They found significant correlations between working memory capacity and reading comprehension, regardless of whether processing in the working memory task involved reading sentences or performing mathematical operations, or whether the items to be remembered were words or digits. However, when mathematical skills (as
measured by scores on the Quantitative Scholastic Aptitude Test) were partialled out of the correlation, only those complex span tasks involving words as the to-be-remembered items remained significant. Turner and Engle (1989) take these results to suggest that it is the domain (i.e., words or digits) of the storage component in complex span tasks that is critical in predicting performance on higher level tasks, while the domain of the processing component does not impact on this relationship. These findings highlight the fact that even working memory measures involving mathematical skills in the processing component can be predictive of verbal ability on the SAT. Turner & Engle also found that the level of difficulty of the processing task is an important factor in ability to predict performance on the Verbal Scholastic Aptitude Test (VSAT); moderately difficult background (i.e., processing) tasks in both reading span and mathematical operations span measures correlated with ability on the Nelson-Denny Reading Test and the VSAT, while more simple and more difficult tasks did not. Thus, the level of difficulty of the task must be taken into consideration in constructing and evaluating working memory measures.

Another critical factor in determining the predictive ability of working memory tests involves the way in which processing on the criterion test is measured. The majority of previous research has employed off-line measures of reading comprehension (e.g., the VSAT or the Nelson-Denny in studies by Turner & Engle, 1989, Engle, Cantor & Carullo, 1992 and Daneman & Carpenter, 1980) and of sentence processing (e.g., sentence acceptability judgements in Waters & Caplan, 1997). Off-line measures may exaggerate individual differences due to additional “post-interpretive” processes which have taken place (Small et al., 2000). Differences in the way working memory and language or sentence processing have been measured may explain some of the conflicting results in previous research (see
Kempler, Almor, Tyler, Andersen & MacDonald, 1998, for an example in which stronger correlations were found between working memory and off-line rather than on-line measures). This in turn has contributed to conflicting views on the nature of the resources being tapped by working memory tests, and the reliability and validity of the tests themselves (e.g., Waters and Caplan, 1996a; Walenski & Swinney, 1999).

Other research has focused on the relationship between performance on working memory tests and performance on other cognitive tasks in different domains. For example, Engle et al. (1992) conducted a study that examined the relationship between the processing and storage components of working memory and its predictive ability for higher level tasks. They used a working memory measure similar to the reading span task. In this task, subjects performed a mathematical operation in an incremental “moving window” method and following each operation were presented with a word to be recalled (e.g., \( \frac{6}{2} - 1 = \_ \_ \_ \) CLOWN). The researchers found no differences between high and low span subjects with respect to processing times on the operation, both in the recall and no recall conditions. There was no correlation between accuracy on the operation alone and the Verbal Scholastic Aptitude Test (VSAT). This led them to conclude that individual differences on working memory tasks do not arise due to differences in processing ability or due to strategic allocation of resources (a trade-off between storage and processing) when information must be recalled. However, it is possible that these null results are due to task complexity; the operations used in this study were relatively simple (e.g., \( \frac{6}{2} - 1 = \_ \_ \_ \)), and therefore they may have failed to differentiate high and low span subjects. As mentioned earlier, the difficulty of the span task determines how well it will predict performance on a criterion task; tasks which are too simple, or too difficult are not as effective in predicting performance on higher level tasks.
Furthermore, there was a significant (albeit weak) correlation of .34 between the number of words recalled and an off-line measure of language comprehension, the VSAT. This demonstrates that a storage component is required to predict ability on a higher level task. It also suggests that working memory is not domain-specific, since in this case performance on a mathematical operations task is able to predict reading ability. Engle et al. conclude that this pattern of results supports the idea of a general working memory capacity, under which working memory is a general, stable characteristic, and thus its predictive ability for higher level tasks does not change when measured by reading or mathematical operations. They claim that individuals vary in the amount of total activation available, which leads to individual differences (cf. Just & Carpenter, 1992). The authors claim that an essential component of their working memory span task is that it "measures the amount of activation available to the individual on a moment-to-moment basis" because of the attention-switching required between processing and storage, and because of the on-line nature of the task (Engle et al., 1992, p.991). Thus, attention-switching between processing and storage functions is the important factor in the predictive ability of working memory tasks, while domain of the processing task is not seen as an important predictive factor.

The operations span task used in the Engle et al. (1992) study addresses a limitation of the reading span task as a measure of working memory: the task itself (reading sentences) is highly similar to what it is meant to predict (reading ability) (cf. Daneman & Carpenter, 1980). Thus, a correlation between reading span and reading ability is expected. Both Turner and Engle (1989) and Engle et al. (1992) provide additional evidence in an operation span task (similar to Daneman and Carpenter's) that the task does not need to involve reading sentences, but that mathematical operations substitute quite well in predicting reading
comprehension. The fact that nonverbal measures, such as mathematical operations, are equally able to predict reading ability (when task complexity is appropriate) has been taken by some researchers as support for the validity of the Daneman and Carpenter reading span measure despite its similarity to the criterion task (Daneman & Merikle, 1996).

While reading span and mathematical operation span working memory tasks have been found to correlate with one another and are able to predict performance on higher-level tasks (e.g., the Nelson-Denny Reading Test, VSAT), researchers have not found similar correlations using other nonverbal tasks such as visuo-spatial tasks. Daneman & Tardif (1987) developed three span tasks with verbal, math or spatial components and compared ability on these tests to reading ability as measured by the Nelson-Denny and vocabulary on the General Aptitude Test Battery. Verbal span and math span both correlated significantly with tests of reading comprehension and vocabulary, whereas spatial span did not. Daneman and Tardif interpret these results as pointing towards a language-specific working memory system in that both math and verbal spans were based on symbolic information. Critics (e.g., Engle et al., 1992) have argued that the correlations involving spatial span must be interpreted with caution because performance on this span task was almost at ceiling, resulting in a limited range of scores. However, other studies, which did not suffer from this limitation, uncovered the same lack of relationship between visuo-spatial working memory tasks and tests of general abilities such as VSAT or paragraph comprehension (e.g., Shah & Miyake, 1996 and Kyllonen, 1993). Although relatively few studies have investigated the relationship between visuo-spatial measures and reading comprehension, the results to this point find little or no correlation between the two. To my knowledge, no study has yet been published which investigates the relationship between visuo-spatial working memory and
sentence processing. It remains to be seen whether more controlled, process-specific measures of visuo-spatial capacity will show a relationship to more process-specific comprehension at the sentence level.

1.9 Summary

The literature reflects a lack of agreement with regards to the nature of working memory and its relationship to sentence processing. Proponents of a general working memory capacity maintain that a single system underlies processing in all domains (Turner & Engle, 1989, Engle et al., 1992). Other researchers have found evidence supporting a division of working memory to support verbal (and possibly mathematical) versus visuo-spatial domains (Daneman & Tardif, 1987, Shah & Miyake, 1996). While some have modelled verbal working memory underlying language comprehension (Just & Carpenter, 1992), others advocate further division of verbal working memory resources for specific linguistic processing (Caplan & Waters, 1999). A more recently developed view on working memory is that experience with language, rather than a separate resource system for storage and processing, can explain individual differences in language processing (MacDonald & Christiansen, in press).

In spite of this controversy, previous research has provided some insight into the relationship between working memory and language processing. In a meta-analysis of the literature published before 1995, Daneman and Merikle (1996) provide ample evidence that working memory tasks are predictive of ability on both general and specific measures of language comprehension. In addition, previous research has helped to identify factors which can influence results. A study by Turner and Engle (1989) underlined the importance of taking task difficulty across domains into account. It was also suggested that one contributing
factor to the conflicting results across studies has been the varying use of on-line versus off-line criterion measures. With respect to sentence comprehension, off-line measures are not sufficiently sensitive to individual differences (Tyler, 1992). Thus, the prevalent use of off-line measures, even for specific language processing measures (see Daneman & Merikle, 1996), is a limitation in past research. Through both findings and methodological limitations, earlier literature has helped to guide the direction of future studies in this area.

1.10 Research Questions and Hypotheses

The present study will address the general question of the nature of the ‘resources’ underlying sentence processing and the extent to which processing in different domains (linguistic, mathematical and visuo-spatial) make use of a similar set of constraints. That is, for a particular set of processing demands, do our minds recruit the same ‘resources’ for tasks in different domains? Or, is the capacity underlying sentence processing unique? This research takes a connectionist stance by arguing that when the nature of processing demands are matched across tasks and domains, working memory tasks from other domains should be equally good, if not better than reading span measures, as predictors of sentence processing ability. The underlying reasoning is that in connectionist models, comprehension results from the recognition of patterns of activation based both on experience and the architecture of the processing network (SRN). Therefore, if tasks in other domains are constructed with sequences or patterns of processing demands over time that are similar to sentences, then they should involve similar patterns of activation or ‘working memory’ demands.

The proposed study will constitute a unique contribution to this field by considering a number of variables that have heretofore not been addressed in one study. First, it will measure the effects of two types of sentence processing demands: syntactic complexity
(object-subject, subject-object) and syntactic ambiguity (main verb, reduced relative).

Second, on-line measures of sentence processing will be used, which more accurately reflect moment-by-moment processing demands. Third, four working memory tasks will be employed that represent different theoretical approaches and cognitive domains (verbal and nonverbal). Significantly, two of these will be new working memory tasks, mathematical and visuo-spatial operations, constructed so as to match the sequence of processing demands found in complex sentences.

Two versions of the mathematical and visuo-spatial operations tasks will be constructed, to mirror the demands in subject-object and object-subject sentences. The subject-object versions will involve an embedded operation within the main task, such that some information will need to be held in memory while other processing takes place, and then the stored information will later need to be integrated to complete the task. This is much like the sequence of processing demands found in subject-object sentences, in which the first noun phrase and the noun phrase of the embedded clause must be held in memory until the first verb is encountered and thematic roles can be assigned. For the mathematical operations task, the subject-object (SO) condition will involve an operation within brackets (e.g., \(9 \div (8 - 5) + 4 = \_\_\) ). The object-subject (OS) version will consist of the same number of operations, but will not involve any embedding (e.g., \(9 \div 3 + 1 - 3 = \_\_\) ).

With respect to the visuo-spatial operations task, participants will view several frames with shapes on them, and will be required to perform an analogy. For the subject-object version, the analogy will be of the form: "A is to D" as "B is to C". Participants will need to recall A while they determine the relationship between B and C (based on shading, location and orientation of the shapes). They will then choose which of two frames (D) completes the
analogy. The object-subject version of this task will involve the same frames, presented in a
different order, such that the analogy becomes "A is to B" as "C is to D". Thus, the object-
subject version of this task does not involve a heavy storage and processing load.

An example of the subject-object version of this task is provided below:

```
A  B  C  D

Screen/Region 1 Screen/Region 2 Screen/Region 3
```

To summarize, the following research questions are addressed in this study:

1. Do working memory tasks in different domains (linguistic, mathematical and visuo-
spatial) tap into a single 'resource'?

2. Is the ability to comprehend syntactically complex or syntactically ambiguous
sentences modulated by 'working memory capacity'?

3. Does the predictive ability of a working memory task rely less on the domain of
processing (e.g., linguistic, mathematical) and more on the similarity of its processing
characteristics to the sentence processing task?

From the above questions, the following research hypotheses have been formed:

1.10.1 Sentence Processing

1. All participants will have greater difficulty (longer response times and more
comprehension errors) with syntactically complex sentences (i.e., subject-object
constructions) than simpler constructions (i.e., object-subject) of equal length. The
difference in response times will be most noticeable at the first and second verb of the
subject-object constructions, which are hypothesized to carry the greatest processing load (cf. King & Just, 1991; Caplan & Waters, 1999).

Reasoning: The subject-object constructions are less frequently encountered, and take longer to process since they are left-branching and have noncanonical ordering of thematic role assignment. Previous research has demonstrated that these dimensions are important in determining sentence complexity (e.g., King & Just, 1991; Small et al., 2000; Caplan & Waters, 1999).

2. All participants will have greater difficulty (longer response times and more errors) comprehending syntactically ambiguous than unambiguous sentences. The locus of this difference in response times will be at the first point of disambiguation in the sentence. In the case of main verb constructions, this point corresponds to the clause following the main verb clause. For reduced relative sentences, the first point of disambiguation is the main verb clause following the embedded reduced relative clause. Increased response times are also expected at the final word in the ambiguous sentences (the final point of disambiguation) as compared with their unambiguous counterparts.

Reasoning: Ambiguous sentences require the reader to entertain multiple hypotheses of syntactic structure and/or exploit other lexical constraints which place demands on processing resources (cf. MacDonald et al., 1992; Kemtes & Kemper, 1997).

1.10.2 Working Memory

3. All participants will have more difficulty with the more complex (subject-object) version of math and visuo-spatial operations than with the less complex (object-subject) version of each task.
Reasoning: For the mathematical operations working memory task, the ‘SO’ version is expected to carry the greatest processing load, since the participant must complete the operation within the brackets while maintaining the first number and operation of the equation, and finally use the result from within the brackets to carry out the operation. This is compared with the ‘OS’ version, in which the participant completes the same number of operations. However in this version the operations are presented sequentially, much like in an object-subject sentence and therefore storage of partial results while computing a second operation is not required.

Similarly, in the visuo-spatial operations task, the ‘SO’ version is expected to carry the greatest processing load, because the participant must find the relationship between the two frames presented, while retaining the visuo-spatial information from the first frame for later integration. This is compared to the ‘OS’ version, in which the participant must also determine the relationship between two frames, but without any storage load.

4. The strength of correlations will vary for performance across the four working memory tasks: Daneman and Carpenter’s reading span, Waters and Caplan’s reading span (span and composite), mathematical operations and visuo-spatial operations.

Reasoning: To the extent that these tasks rely on overlapping neural networks (i.e., similar sequential processing characteristics), they will be correlated. Since the latter two tasks are designed to mirror the processing demands in complex sentences, they should be the most strongly correlated.
1.10.3 Sentence Processing and Working Memory

5. Correlations varying in strength will be observed between performance on the sentence processing task and each working memory task.

a) Reading span: Performance on both the Daneman-Carpenter reading span task and the Waters-Caplan reading span will significantly correlate with performance in syntactic complexity and ambiguity conditions.

Reasoning: Both reading span and sentence processing measure language processing skills, i.e., the reading span tasks require a subset of skills that overlaps with the skills required in processing syntactically complex and ambiguous sentences (cf. MacDonald & Christiansen, in press).

b) Mathematical operations: Performance on the mathematical operations working memory task will significantly correlate with performance on syntactically complex but not ambiguous sentences.

Reasoning: Because the task is designed to match the sequence of processing demands in syntactically complex sentences, it should draw on similar patterns of activation as those found in processing these sentences. On the other hand, because the sequence of processing demands is different for syntactically ambiguous sentences, correlations with mathematical operations should be weak. This finding would also support the capacity theory, in that mathematical operations have been considered by Just and Carpenter (1992) to involve symbolic operations, and thus overlap with language processing. Caplan and Waters' view on working memory for sentence processing does not predict a correlation here, because these authors consider sentence
processing ability as separate from the ability to perform other verbal (e.g.,
mathematical) tasks.
c) Visuo-spatial operations: Performance on the visuo-spatial operations working
memory task will correlate with processing of syntactically complex but not
ambiguous sentences.
Reasoning: As with the mathematical operations task, this task was created to
match the moment-to-moment processing demands in the sentence processing
task, and therefore, from a connectionist perspective, should draw on the
subset of skills involved in processing syntactically complex but not
ambiguous sentences. The capacity theory would not predict a correlation
between performance on visuo-spatial operations and sentence processing, nor
would the separate sentence interpretation resource theory, as both predict that
nonverbal (e.g., visuo-spatial) tasks draw on a different pool of resources from
that supporting sentence processing.
6. Performance on the sentence processing tasks will be predicted in varying degrees by
each working memory task. The working memory tests created by the author
(mathematical operations and visuo-spatial operations) are hypothesized to account
for a greater amount of variance in performance on complex sentences than are the
reading span tasks.
Reasoning: Both of these working memory tasks were created to specifically match
the sequence of processing and storage demands found in complex sentences. From a
connectionist point of view, since the patterns of activation required are similar,
performance on these tasks will produce the strongest correlations.
CHAPTER 2: METHOD

2.1 Overview

The following chapter details the experimental variables and conditions employed in this study. The construction of stimuli in the sentence processing and working memory tasks is explained in detail, along with the procedure for each task. Finally, the design used for data analysis is outlined.

2.2 Participants

Thirty-one participants between the ages of 18 and 30 were recruited from the University of British Columbia and the surrounding community. Data from one participant was excluded from the study, due to technical problems during testing. Ten of the participants were male, 21 were female. Participants in the study had an average age of 24.7 years (SD = 3.3, range = 19 to 30 years) and an average of 16.9 years of education (SD = 2.4, range = 14 to 24 years). All participants were native speakers of English and were determined through a vision screening to have normal or corrected-to-normal vision (i.e., using a Snellen eye chart, better than 20/30 vision at a test distance of ten feet). Participants received $20.00 for their participation.

2.3 Materials and Procedure

2.3.1 General Procedure

Participants were tested individually in a single session that lasted 1 1/2 to 2 hours. Participants’ vision was screened before beginning the experimental tasks. Vision in both eyes was tested using a Snellen eye chart at a test distance of ten feet. Following standard optometrical procedure, participants were asked to read a series of five letters, starting at the 20/20 line. If they identified 4/5 letters correctly, the screening ended and they were assigned
a rating of 20/20 vision. If more than one mistake was made, screening continued with the next largest letters (20/25), and so on until they correctly identified 4/5 letters at a level. The experimental tasks were presented on a 21-inch computer monitor, with a resolution of 1024 by 768 pixels. Participants were seated at a comfortable distance from the monitor (approximately 2 feet away). The E-Prime program Beta 4.0 version (Psychology Software Tools, Inc., 2000) was used to control the presentation of stimuli in all tasks and to record the accuracy of participants’ responses and their response times.

All participants completed the criterion sentence processing task first and then proceeded to the working memory tasks. The order of the four working memory tasks (Daneman-Carpenter reading span, Waters-Caplan reading span, mathematical operations and visuo-spatial operations) was counterbalanced across participants in a Latin square design. The procedure for each of the experimental tasks will be further described below.

2.3.2 Measure of Sentence Processing

An on-line measure of sentence processing was employed in this study. The use of an on-line measure permits a view of the demands of sentence processing as it unfolds over time. Off-line measures of sentence processing, such as response time for sentence acceptability judgements (Waters & Caplan, 1997), sentence-picture matching (Small et al., 1997), or sentence enactment (Caplan & Waters, 1999) capture end-of-sentence comprehension once all computations have taken place, and are consequently more contaminated by memory and attentional demands and strategies. This poses a serious limitation, since sentence processing is normally an unconscious and automatic event (Tyler, 1992). Tyler has shown that viewing participants’ performance as it unfolds on-line may uncover individual differences which are not apparent when using off-line measures.
The moving window method (Just, Carpenter & Woolley, 1982) was chosen as the on-line measure of sentence processing as this has been a commonly used technique in previous on-line studies of individual differences in language processing (e.g., King & Just, 1991, MacDonald, Just & Carpenter, 1992 and Daneman & Carpenter, 1983). Furthermore, moving window methods have been evaluated as “comparable in their resolution” to eye-tracking procedures (i.e., similar sensitivity to moment-by-moment effects of syntactic complexity and ambiguity) (Haberlandt, 1994). In the moving window method, a portion of text (usually a sentence) is displayed on a computer monitor, with dashes replacing the letters of each word. The participant presses a button to reveal the first word and continues to press the button to reveal subsequent portions of the sentence, one word at a time. As each successive word is revealed, the previous word returns to dashes, so that readers cannot return to words they have already read. This gives the impression of a window moving across the computer screen. Some researchers (Kemtes & Kemper, 1997) have omitted the dashes in the moving window method so as to minimize the effects of ‘parafoveal processing’ of previously read and to-be-read words in the sentence. By removing cues to word and sentence length, one can more accurately measure local effects of syntactic ambiguity or complexity. In moving window studies, the time between button presses is taken as a measure of the processing time required for that segment of the sentence. Thus, longer response times to certain segments of a sentence indicate that the processing load is greater for that segment. This approach assumes that the computational demands of storage and processing can be measured separately. However, at every point in the sentence, there are concurrent demands on both storage and processing. Therefore, in this study, the dependent measure of response times is assumed to reflect both storage and processing. Comprehension questions follow each
sentence to ensure that participants were indeed reading and processing the sentences for accurate comprehension.

2.3.3 Sentence Complexity Variables

2.3.3.1 Introduction

Two sentence processing variables were used as criterion measures in this study: syntactic complexity and syntactic ambiguity. These variables were chosen because clear predictions of performance can be made based on the known processing demands of the two types of sentences. The syntactic complexity variable consisted of two conditions, subject-object (SO) and object-subject (OS) sentences, in a replication of Caplan and Waters (1999) and a partial replication of King and Just (1991). The syntactic ambiguity condition involved two types of sentences with embedded relative clauses, one in which a temporary ambiguity was introduced at the first verb, and the other which had no ambiguity (cf. MacDonald, Just and Carpenter, 1992; Pearlmutter & MacDonald, 1995).

2.3.3.2 Syntactic Complexity

Centre-embedded relative clause sentences have been cited by King and Just (1991) as the “classic example” of a syntactic structure that places heavy demands on working memory. The most demanding centre-embedded clause is that in which the matrix clause head NP is the object of the relative clause, for example: *The officer, who the clerk thanked, walked to the door.* The right-branching object-subject structure, as in *The officer thanked the clerk, who, walked to the door,* is more easily comprehended. The differences in the demands placed by the subject-object (SO) and the object-subject (OS) can be explained in terms of processing load. These sentence types differ with respect to two dimensions: canonicity of thematic role mapping and branching direction. Canonicity of thematic role
mapping refers to whether assignment of ‘agent’ and ‘theme’ roles proceeds in the sequence most typical in English: Noun 1 (agent)-Verb-Noun 2 (theme). Branching direction refers to whether an embedded relative clause branches from the subject noun phrase (left-branching) or the object noun phrase (right-branching). In the right-branching OS construction, assignment of thematic roles within the matrix and embedded clause proceeds in canonical order; the first NP (the officer) is assigned the role ‘agent’ by the first verb encountered, and the second NP (the clerk) is the ‘patient’. Thus, information need only be retained over a short distance before roles can be assigned. In the SO construction on the other hand, assignment of thematic roles proceeds in non-canonical order. Furthermore, this is a left-branching construction (i.e., the embedded relative clause branches from the subject) which requires the first NP to be retained in memory across the embedded relative clause until the main verb of the sentence is encountered and thematic roles can be assigned. According to Gibson (1998) there is a “memory cost” for incomplete syntactic dependencies related to thematic role assignment, such that NPs that have not yet been assigned thematic roles, and verbs that have not yet found arguments, place demands on storage. In the above SO sentence, the main clause (The officer walked to the door) is interrupted by the embedded clause (who the clerk thanked), necessitating storage of the first NP (the officer) for later thematic role assignment. Thus, there is an increase in working memory demands as the first verb of the embedded clause (thanked) is encountered, since two thematic roles must be assigned (i.e., the role ‘patient’ is assigned to officer, and the role ‘agent’ is assigned to clerk). Based on the sequence of demands in SO sentences, it is predicted that the two adjacent verbs will be slower to process than other parts of the sentence and will also result in longer and less accurate responses to comprehension questions (cf. King & Just, 1991).
The prediction made by King and Just, and also in this study, is that all subjects will have more difficulty processing SO sentences as opposed to OS sentences. However, those who perform worse on the working memory tasks will have substantially more difficulty with SO sentences than those with higher working memory scores, because they have less experience with language, and therefore have had less exposure to this less frequent structure. Thus, it is hypothesized that performance on complex sentences will correlate with performance on working memory tasks.

2.3.3.3 Materials

In the syntactic complexity condition, a replication of Caplan and Waters (1999) and a partial replication of King and Just (1991) was carried out. However, unlike the King and Just study, which used subject-subject and subject-object constructions, the stimuli in this study consisted of subject-object (SO) and object-subject (OS) sentences. Initially, a full replication of Caplan and Waters (1999) was planned. However, the sentences used in that study contained a great deal of overlap in lexical content. Furthermore, the sentences contained a mix of animate and inanimate noun phrases, and were therefore not reversible. For these reasons, a new set of sentence stimuli was created for use in this study. The sentences were constructed such that the first two noun phrases were animate and semantically reversible. Subject-object and object-subject versions of each sentence were created. For example:

1. *The detective who the woman believed reported the evidence.* (SO)

2. *The detective believed the woman who reported the evidence.* (OS)

An effort was made not to repeat any content (i.e., noun phrases, verb phrases) across sentences. Sentences ranged in length from nine to eleven words, with an average length of
9.76 words per sentence. A total of 100 sentences were constructed, and divided into two lists of 50 sentences, such that half the sentences were SO and half were OS constructions, with no overlapping content within a list. Each participant saw sentences from only one of the two lists. A complete copy of the SO and OS sentences appears in Appendix A.

A set of four comprehension questions was written for each sentence. Comprehension questions were written in yes/no form, questioning both the embedded and matrix clause of the sentence. A yes and a no response version was written for each question. For example, for the object-subject construction ‘The teacher admired the student who won the award’ the questions were:

Matrix Clause

Did the teacher admire the student? (yes)
Did the student admire the teacher? (no)

Embedded Clause

Did the student win the award? (yes)
Did the teacher win the award? (no)

For the subject-object version of this sentence, The teacher who the student admired won the award, the questions were as follows:

Matrix Clause

Did the teacher win the award? (yes)
Did the student win the award? (no)

Embedded Clause

Did the student admire the teacher? (yes)
Did the teacher admire the student? (no)
The four lists of questions were combined with the two sentence lists to create a total of eight lists. Each subject saw only one question per sentence, and was presented with equal numbers of each question type (embedded and matrix). For each list, the correct answer for half the questions was 'yes'.

2.3.3.4 Syntactic Ambiguity

Previous research has provided detailed accounts of the sequence of demands in ambiguous sentences (e.g., Tanenhaus & Trueswell, 1995). The type of ambiguity employed in this study was main verb/reduced relative clause ambiguity (MV/RR), in a partial replication of the MacDonald et al. (1992) experiment. An example taken from that experiment is:

1. The soldiers warned about the dangers...before the midnight raid.
2. ...conducted the midnight raid.

In these sentences, the verb warned can be either the main verb of the sentence (as in the first example), which would assign the role 'agent' to soldiers, or it may introduce a reduced relative clause (as in the second example), in which case soldiers would be the 'patient' of the embedded verb (warned).

MacDonald et al. divided sentences into three regions for analysis:

2. *The experienced soldiers [warned about the dangers]$_1$ [conducted the midnight]$_2$ [raid]$_3$.*

For sentences with main verb interpretation, response times for all subjects are expected to increase over the second region of the sentence (e.g., *before the midnight*) in which the
ambiguity is maintained, and will be longest on the final word (raid), which is the point of disambiguation. In the case of sentences that take the reduced relative interpretation, the second region of the sentence, which introduces the main verb (e.g., conducted the midnight) is the first point of disambiguation, and is expected to show longer response times than the first region. The final word (raid) is also expected to show elevated response times, as the final point of disambiguation.

2.3.3.5 Materials

The syntactic ambiguity condition was a partial replication of the MacDonald et al. (1992) study and the Waters and Caplan (1996b) study. Sentences in this condition were main verb/reduced relative constructions, half of which were unambiguous and half of which were temporarily ambiguous. The stimuli from the MacDonald et al. (1992) study and those of Kemtes and Kemper (1997) were used for this sentence processing variable. There were four sentences in each set. Two of the sentences were temporarily ambiguous, one resolving in favour of a main verb interpretation, the other resolving in favour of a reduced relative reading. Each of these sentences was paired with an unambiguous counterpart. For example:

Main verb -- Ambiguous

The soldiers warned about the dangers before the midnight raid.

Reduced relative -- Ambiguous

The soldiers warned about the dangers conducted the midnight raid.

Main verb -- Unambiguous

The soldiers spoke about the dangers before the midnight raid.

Relative clause – Unambiguous

The soldiers who were told about the dangers conducted the midnight raid.
The original MacDonald et al. stimuli incorporated eight verb triplets, such as *spoke, told* and *warned*. The first verb (e.g., *spoke*) allowed only a main verb interpretation of the sentence. The second verb (e.g., *told*) permitted only a relative clause reading of the sentence. Thus, two of the verbs in the triplet lead to unambiguous sentence resolutions. The third verb (e.g., *warned*) introduced a temporary ambiguity between a main verb and a reduced relative reading of the sentence. MacDonald et al. reported that the verbs in each triplet did not differ significantly in length or in frequency. The authors stated that sentence frames were created such that:

For each triple, three unrelated sentences were written, producing 24 distinct sentence frames. In most cases, the three verbs in a triplet were semantically related. In some cases, they were less related, but an effort was made to keep all three verbs equally compatible with the sentence frames in which they occurred (MacDonald et al., 1992, p. 63).

Kemtes and Kemper (1997) added another four verb triplets to the original stimuli set, increasing the total number of sentence stimuli from 96 to 144 sentences. More recently, another four verb triplets were added (Kemtes, 2000), creating a total of 16 verb triplets and increasing the total number of sentences to 192. These sentences ranged in length from 10 to 14 words, with an average length of 11.58 words per sentence.

The total set of 192 sentences was divided into four lists according to the procedure followed in previous studies (MacDonald et al., 1992; Kemtes & Kemper, 1997). Each list consisted of a total of 48 sentences, with 12 of each sentence type (main verb ambiguous, reduced relative ambiguous, main verb unambiguous, relative clause unambiguous). The content (i.e., subject and object) particular to a sentence frame was not repeated within a list.
Thus, in the case of the exemplar sentences listed above, one of these sentences was assigned to each list. Sentences were assigned to lists such that 8 of the 16 ambiguous verbs appeared with both main verb and reduced relative resolutions in a single list. The remaining eight ambiguous verbs appeared on a list in only one ambiguous sentence type, that is, four occurred in sentences with main verb resolution, and four in sentences with reduced relative resolution.

Each sentence was paired with a comprehension question, which was written with both ‘yes’ and ‘no’ response versions. The original questions were altered slightly, so as to eliminate any ambiguity in the questions themselves. A complete list of the sentence stimuli and comprehension questions appears in Appendix B.

2.3.3.6 Fillers

A total of 52 filler sentences were included, so as not to over-sensitize subjects to the experimental sentence types. The syntactic complexity (OS/SO) and syntactic ambiguity (MV/RR) variables also served as fillers for each other. The filler sentences were a variety of sentence types (e.g., right-branching, left-branching and simple sentences), a subset of which were stimuli for another experiment. Filler sentences varied in length proportionately to the sentence lengths of the experimental sentences.

2.3.3.7 Procedure

Sentence stimuli for the syntactic complexity and syntactic ambiguity conditions, along with fillers, were combined into one experiment, yielding a total of 150 sentences per list. The sentence processing experiment consisted of eight lists, counterbalanced for sentence content and question type. For each list, the correct answer to half the questions was ‘yes’. The presentation of these lists was counterbalanced across participants.
The sentences were presented using the moving window technique (Just, Carpenter & Woolley, 1982). All sentences began with an asterisk as a fixation point. By pressing the spacebar, participants were able to advance the screen to view each successive word in the sentence. As a new word appeared on the screen, the previous word disappeared. Unlike in earlier moving window studies (e.g., MacDonald et al., 1992; King & Just, 1991), dashes were not used to replace the words on the screen. The dashes were omitted to minimize parafoveal processing of words surrounding the to-be-read word (cf. Kemtes & Kemper, 1997). Words were presented in Courier New 14 point font, and appeared in black type on a white background.

Participants were instructed to read the sentences as quickly as they could, but to ensure that they understood what they were reading. It was emphasized that they try to understand the sentences as they read them, so that they could answer the questions as accurately as possible. Participants were presented with five practice trials, which had to be completed at an 80% accuracy level before they could proceed to the experimental trials. Sentences were presented in random order. Each sentence was followed by a yes/no comprehension question. Although feedback on accuracy was given during the practice block, participants did not receive feedback on their performance during the experimental trials. Following the first 75 sentences, participants were permitted to take a short break before completing the second half of the sentences.

2.3.4 Working Memory

There is a lack of consensus among researchers with regards to how working memory should be measured. Two versions of the reading span task, which was first proposed by Daneman and Carpenter (1980), were used in this study. In the interests of replicating
previous work, the original Daneman and Carpenter task was employed. Daneman and Carpenter's task includes both processing and storage components, but the dependent measure is the number of words recalled. No direct measure of processing accuracy is included in this task. Waters and Caplan (1996a) claim that Daneman and Carpenter's reading span task is not a stable measure of working memory. They found that 41% of subjects changed span categories (low, medium or high) when re-tested on the Daneman-Carpenter reading span test three months after the initial test. Moreover, 18% of subjects changed between more than one category (i.e., high span became low span, or vice versa). Waters, Caplan and Hildebrandt (1987) and Waters and Caplan (1996a) developed another reading span task which incorporates explicit measures of both processing accuracy and storage (on four sentence types: cleft subject, cleft object, object-subject, and subject-object). Scores on this task were expected to be more stable, given that all aspects of performance on the task were measured, taking into account possible trade-offs in storage and processing. This task was found to be more reliable when participants were re-tested after a three month interval (Waters & Caplan, 1996a). For this reason, the Waters and Caplan reading span task was used in this study as a second measure of working memory capacity. Only the cleft subject sentences version of the task was administered since the other types overlap with the sentence types used in the sentence processing criterion measure. This version has also been used by Waters and Caplan in later studies that addressed individual differences in sentence processing (1996b). One drawback to both of these language-based working memory tasks is the degree of similarity to the criterion measure of sentence processing. That is, both tasks involve reading sentences. The present study therefore introduces two other working memory tasks not requiring sentence processing.
The two new working memory tasks were designed to require mathematical operations and visuo-spatial processing. These tasks were constructed so as to be carefully matched to the syntactic complexity variable of the sentence processing task in terms of difficulty and the sequence of processing and memory demands as the task unfolds. The tasks were constructed to contain storage demands intrinsic to the task itself rather than explicit recall; storage and processing demands were analogous to those in sentence processing in that information needed to be stored for later integration in the task. In contrast to traditional span tasks, but similar to the sentence processing task, there was no measure of storage independent of the task itself. Instead, tasks were administered in a moving window method, where response times reflected the demands of both concurrent storage and processing.

2.3.4.1 Daneman and Carpenter’s Reading Span Task

Stimuli for this task were taken from the original Daneman and Carpenter reading span task (1980). The test consisted of 100 unrelated sentences, divided into five levels, ranging from 13 to 16 words in length. For example, one set at the two-sentence level was:

_I was so surprised at this unaccountable apparition, that I was speechless for a while._

_When at last his eyes opened, there was no gleam of triumph, no shade of anger._

None of the sentence-final words (to-be-remembered words) were repeated in this test. To-be-remembered words were nouns, verbs, adjectives, adverbs and pronouns. The entire set of sentences is listed in Appendix C.

Participants completed three practice trials at the two-sentence level before proceeding to the test trials. Participants were instructed to read the sentences aloud at their own pace, and to remember the last word of each sentence. They were told to proceed to the next sentence immediately after they finished reading. The test sentences were organized into blocks of 2,
3, 4, 5 and 6 sentences, with five sets for each level. At the end of a set, participants were prompted with an asterisk to repeat back all the sentence-final words they could remember. They were asked to recall the words in order, if possible, and not to say the last word first, unless it was the only word they could remember. The examiner warned participants each time before the set size increased. Participants' responses were recorded by the examiner. Testing was terminated when a participant failed 4/5 sets at a particular level.

Administration of this reading span task followed the general procedures outlined by Daneman and Carpenter (1980). Sentences appeared in their entirety on a computer screen. They were presented in Courier New 14 point font, with black type on a white background. All participants read the same sentences, in the same order.

2.3.4.2 Waters and Caplan's Reading Span Task

The second reading span test was that developed by Waters and Caplan (1996a), using a subset of materials from an earlier study (Waters, Caplan & Hildebrandt, 1987). Stimuli for this task were obtained from Waters and Caplan (1996a; 1996b). Only the cleft subject sentences version of the reading span task was administered in this study (cf. Waters & Caplan, 1996b). It consisted of 100 unrelated sentences, all of which were cleft subject constructions. The sentences ranged from 8 to 11 words in length, with a mean number of 8.7 words for acceptable sentences, and 8.1 words for unacceptable sentences. The sentences were constructed such that:

Half of the sentences of each type had verbs which require animate subjects and inanimate objects (e.g. "It was the man that clenched the pillow.") and half had verbs which require inanimate subjects and animate objects (e.g. "It was the toy that fascinated the child."). The sentences were all semantically irreversible.
Unacceptable sentences were formed by inverting the animacy of the subject and object noun phrases (e.g. "It was the pillow that clenched the man."). This leads to violations of "selectional restrictions," the inherent restriction verbs place on semantic features of their arguments. Recognition of the acceptability or unacceptability of a sentence in this task requires a syntactic analysis, assignment of thematic roles to noun phrases, and comparison of the thematic roles with selectional restrictions of the verb. (Waters, Caplan & Hildebrandt, 1987, p. 538-539.)

Unlike the Daneman and Carpenter reading span task, to-be-remembered words in this task were all nouns, some of which were plural, some singular. There was overlap in the content of the sentences used in this task. The result was that many of the to-be-remembered words were repeated across sets (for example, the word *child* appeared in 10 of the 25 sets, *man* appeared in 7 of the 25 sets). The complete list of sentences used in this task appears in Appendix D.

Also unlike the Daneman and Carpenter reading span task, for this task, participants read the sentences silently and pressed a ‘Y’ or ‘N’ key to indicate if the sentence was acceptable or unacceptable. Immediately following the acceptability judgement, the next sentence in the set appeared on the screen. When the set was complete, a question mark appeared on the screen to indicate that the participant was to recall the sentence-final words by reciting them aloud. Participants were instructed to recall words in the order in which they appeared, if possible, and not to recall the last word first unless it was the only word they could remember. Three practice sets at the two-sentence level were completed before the test trials began.
As in the Daneman and Carpenter task, the sentences were organized into five blocks each of 2, 3, 4, 5 and 6 sentences. There were five sets of sentences in each block. Participants were given a warning each time before the number of sentences in a set increased. A participant's reading span was defined as the highest level at which all sentence-final words were recalled on 3/5 sets (as in the Daneman and Carpenter reading span task). Testing was discontinued at the point at which a participant failed 4 out of 5 sets. The examiner recorded subjects' recall of sentence final words, while accuracy for the sentence acceptability judgements was recorded by the E-prime program (Psychology Software Tools, Inc., 2000).

Presentation of the sentences was identical to the Daneman and Carpenter reading span task; entire sentences were presented on the monitor in Courier New 14 point font, with black type on a white background. Each set in this task began with a fixation point (an asterisk) that appeared in the centre of the screen for 300 milliseconds (cf. Waters & Caplan, 1996a). All participants read the same sentences in the same order of presentation.

2.3.4.3 Mathematical Operations

The mathematical operations working memory task was constructed by the author, in line with the tasks developed by Turner and Engle (1989) and Engle et al. (1992). Two conditions were constructed so as to match as closely as possible the sequence of processing demands in subject-object and object-subject sentences. In the first condition (subject-object), equations such as $5 \times (1 + 3) - 9 = \_\_\_\_$, were generated according to the following constraints:

i. The first number was from 1-9 (for multiplication) or 40-60 (for division), followed by a multiplication or division sign.
ii. The operation inside the brackets was an addition or subtraction of numbers from 1-9. The result of this operation was a number from 1-9.

iii. The result of the first operation (outside the brackets) was a number from 10-20.

iv. The final operation was an addition or subtraction of a number from 1-9. The final answer was a number from 10-20.

Given the format of presentation (i.e., moving window) and the nature of the operation, this task closely resembles processing of centre-embedded clause structures in terms of processing demands. Participants first encounter a number and operation (multiplication or division) to be carried out. However, the open brackets delay the completion of the first operation, requiring participants to retain it in memory (since it will disappear when the spacebar is pressed). Processing inside the brackets must be carried out, and the result of this operation must be integrated with the first operation presented. This parallels the manner in which the centre-embedded clause interrupts the matrix clause, necessitating storage and later integration of that information.

Equations in the second condition were generated to mirror the sequence of demands in object-subject sentences. These equations were generated according to the following guidelines:

i. The first number was from 1-9 (for multiplication) or 40-60 (for division), followed by a multiplication or division sign.

ii. The next number was from 1-9. The response from this first operation was a number from 10-20.

iii. The next operation was an addition or subtraction of a number from 1-9, yielding a result between 10 and 20.
iv. The final operation was an addition or subtraction of a number from 1-9, yielding a final result between 10 and 20.

EXAMPLE: 50 ÷ 5 + 4 + 5 = ___

A total of 40 experimental items was constructed; 20 were operational sequences of the first type described, mimicking the processing demands of subject-object sentences, and 20 mirrored the demands of object-subject sentences. The sequence of operations for each item was counterbalanced within each condition ('object-subject' and 'subject-object' operations). That is, for the 20 operations of each type, 10 items began with a multiplication and 10 began with a division. Each of these sets of 10 was then further divided, such that five continued with an addition, and five continued with a subtraction. Finally, each of these sets of five items were further subdivided. For the items beginning with a multiplication, the final operation was an addition for three of the items, and a subtraction for two of the items. The opposite subdivision was made for the items beginning with a division (i.e., the final operation was an addition for two items, and a subtraction for three items). An effort was made to ensure that equations were balanced with respect to the number of times each number and result occurred within each condition. A complete list of the stimuli used in the mathematical operations task is in Appendix E.

The mathematical operations task was presented in a moving window method (cf. Engle et al., 1992). Participants pressed the spacebar to advance the moving window, and entered their response on the number pad on the keyboard. As soon as one trial (i.e., sequence of operations) was completed, the next trial began with a 'button icon' as a fixation point on the screen. An asterisk was not used as a fixation point, because of possible confusion with its common use as a multiplication sign. The numbers and operations were presented in Courier
Operations were presented in eight sections:

i. First number

ii. Multiplication or division sign

iii. Open bracket and number

iv. Addition or subtraction sign

v. Number and closed bracket

vi. Addition or subtraction sign

vii. Last number and equal sign

viii. Question mark to indicate participant was to enter his/her response

Operations in the 'object-subject' condition were presented in the same manner, except that there were no brackets in these items.

Participants were instructed to advance through the equations as quickly as they could, while ensuring that they performed the calculations as accurately as possible. Participants completed a set of six practice equations before scoring began. A criterion level of performance of 80% was required on the practice items before a participant could move on to the experimental trials. Feedback on performance was provided during the practice block, but was not continued for the experimental trials. Items for the 'subject-object' and 'object-subject' conditions were combined and presented in random order. Participants' responses and response times were recorded by the computer.

2.3.4.4 Visuo-spatial Operations

The visuo-spatial task was developed by the author. Half the items mimicked the sequence of demands of a centre-embedded clause construction, while the other half matched
the demands of simpler object-subject constructions. Both conditions consisted of five frames depicting various shapes. The goal of the task was to determine how two of the frames were the same, based on three possible dimensions, and to create another pair that matched on this same dimension.

In the ‘subject-object’ condition, participants saw one frame, the content of which needed to be retained in memory. Participants were then presented with two frames together, and determined the relationship between these two frames based on one of three possible dimensions: orientation, shading or location of the shapes in each frame (see example below). The next (and last) screen revealed another two frames. The participant then had to choose which of these frames matched the first frame presented, using the relationship determined between the frames on the second screen. This was considered to match the processing demands in centre-embedded constructions in that subjects had to retain information in memory while computing the relationship between intervening information, and then return to integrate the first frame into the operation.

EXAMPLE:

In this example, the first frame must be retained in memory, without knowing what the relevant dimension may be. The relationship between Frames 2 and 3 is location (i.e., the ovals are both located in the two rightmost quadrants of the frames). The frame which fulfills
this relationship (of location) with Frame 1 is Frame 5; both have ovals appearing in the leftmost two quadrants of the frame. Frame 4 does not match Frame 1, since the ovals on this frame appear on the bottom, rather than the left of the frame. Further examples for both conditions using the other dimensions are shown in Appendix F.

The other half of the stimuli were constructed to mirror the demands in the less complex object-subject sentences. The task was essentially the same, however, participants were not asked to maintain the information from the first frame in memory. Instead, participants were shown two frames on the first screen, and asked to determine the relationship between the first two frames presented. They were then presented with a third frame, and were asked to retain this in memory. This time, they only needed to remember the relevant dimension (location, orientation or shading) determined from the first two frames. They then used this relationship to choose from among the final two frames the one which fulfilled this same relationship with the third frame. Examples of the object-subject stimuli are shown in Appendix F.

Items on this task were constructed so that all three dimensions (orientation, location and shading) needed to be considered before making a decision on each trial. Orientation and location had four possibilities, while shading had two possibilities. On the dimension of orientation, shapes appeared either vertically, horizontally, or diagonally (two diagonals). Location of shapes varied in terms of where shapes appeared in two of four quadrants; they could appear in the topmost two quadrants, the bottom two quadrants, the leftmost quadrants (vertically) and the rightmost quadrants. Shading of shapes varied in two ways: shapes were filled or unfilled.
A total of 40 items were constructed; the order of presentation of the frames was counterbalanced across two conditions, so that the stimuli used in the ‘object-subject’ condition for the first participant were the stimuli for the ‘subject-object’ condition for the second participant. The stimuli were constructed so that all possibilities under each dimension were equally represented.

The visuo-spatial task was presented using the moving window technique. In the ‘subject-object’ condition, participants saw a single frame on the first screen in the trial. Participants were instructed to maintain the frame as accurately as possible in memory. The first frame disappeared when the spacebar was pressed, and the next two frames appeared on the screen together. Once participants determined how these two frames were the same, they pressed the spacebar again, to reveal another two frames. The participant then made a choice between frames, and the next trial began. In the ‘object-subject’ condition, the first two screens were reversed in order of presentation. That is, participants were presented with two frames on the first screen, and one frame on the second screen. As in the ‘subject-object’ condition, participants saw two frames on the final screen, where they were required to make a decision.

For both conditions, participants were shown exemplars of the frames on paper before proceeding to the practice trials. The three dimensions, shading, orientation and location of the shapes, were outlined. Next, the order of the frames was explained. Since the ‘object-subject’ condition was the simpler of the two conditions, participants were first instructed in this condition, and completed six practice trials. Participants were required to achieve a criterion level of performance of 80% on the practice trials in order to move on. Once this practice set was completed, the ‘subject-object’ condition was explained to participants, and
another six practice trials followed. Again, a criterion level of 80% was required before moving on to the experimental trials. Feedback on performance was provided during the practice trials, but was not continued for the experimental trials. Participants were instructed to proceed through the items as quickly as they could, while ensuring that their decisions were as accurate as possible.

The 'subject-object' and 'object-subject' conditions were presented in two separate blocks. The order of presentation of these blocks was counterbalanced across participants. Participants' accuracy and response times were recorded by the computer.

2.4 Analysis

2.4.1 Syntactic Complexity

The intervals between button presses were recorded by the E-prime program with millisecond accuracy. This response time was taken as a measure of storage and processing for each portion of the sentence. Although response time was collected for each individual word in the sentence, words were grouped into phrases for analysis (cf. Caplan & Waters, 1999). Thus, sentences in the two conditions were analyzed in the following manner:

Object-subject sentence

\[
[\text{The officer}_{NP1} \text{ thanked}_{V1} \text{ the clerk}_{NP2} \text{ who } \text{ walked}_{V2} \text{ to the door.}]_{\text{Final Phrase}}
\]

Subject-object sentence

\[
[\text{The officer}_{NP1} \text{ who } \text{ the clerk}_{NP2} \text{ thanked}_{V1} \text{ walked}_{V2} \text{ to the door.}]_{\text{Final Phrase}}
\]

This analysis permitted a comparison across object-subject and subject-object conditions. The particular area of interest was the difference between the OS and SO conditions at the second verb (V2). At this point in the sentence the SO version is expected to carry heavier storage and processing requirements than the OS version, since this is the point with the
greatest memory load (owing to its non-canonical order and right-branching structure leaving thematic roles which have yet to be assigned).

2.4.2 Syntactic Ambiguity

One dependent variable for processing syntactically ambiguous sentences was response time. Response times were gathered word-by-word as the experiment unfolded. For the purposes of analysis, sentences were divided into three regions and the mean response time for each region was calculated (cf. MacDonald et al., 1992; Kemtes & Kemper, 1997). Examples of the regions across the four sentence types are as follows:

Main Verb – Temporarily Ambiguous

The experienced soldiers [warned about the dangers]₁ [before the midnight]₂ [raid.]₃

Main Verb – Unambiguous

The experienced soldiers [spoke about the dangers]₁ [before the midnight]₂ [raid.]₃

Reduced Relative – Temporarily Ambiguous

The experienced soldiers [warned about the dangers]₁ [conducted the midnight]₂ [raid.]₃

Relative Clause – Unambiguous

The experienced soldiers who were [told about the dangers]₁ [conducted the midnight]₂ [raid.]₃

As described by Kemtes and Kemper:

The first region corresponds to the introduction of the matrix verb (ambiguous or unambiguous) [this is true for the main verb sentences, but is the relative clause in the reduced relative sentences]; the second region corresponds to the continuation of the clause or phrase and is the first possible point of disambiguation for the reduced
relative clause sentences; and the third region corresponds to the final word of the sentence and is the point of disambiguation for the main verb ambiguous sentences and is the final point of disambiguation for the reduced relative clause sentences (1997, p. 366).

Because the main verb and relative clause sentences constitute distinct constructions, they were analyzed separately. The main areas of interest were differences across ambiguous and unambiguous sentences of each type (main verb and relative clause) in regions 2 and 3.

2.4.3 Daneman and Carpenter's Reading Span Task

A participant's reading span was defined as the highest level at which all sentence-final words were recalled in 3/5 sets. Half credit was given for recall of 2/5 sets at a level. For example, if a participant correctly recalled all sets at level three, but only two sets at level four, that participant would be assigned a reading span score of 3.5.

2.4.4 Waters and Caplan's Reading Span Task

Two scores were derived from performance on Waters and Caplan's reading span task: a recall-only reading span score and a composite z-score. Scores for the former were determined in the same way as for Daneman and Carpenter's reading span; the highest level at which a participant could correctly recall all sentence-final words on 3/5 sets determined that participant's reading span level. Half credit was given for recall of 2/5 sets at a given level. Scores for the Waters and Caplan composite were based on the number of words recalled, as well as accuracy of sensibility judgements and response time for sensibility judgements. Following the procedure outlined in Waters and Caplan (1996a), performance on each of the above measures was first converted to a z-score, and the resulting three z-scores were averaged to obtain each participant's composite working memory score.
2.4.5 New Working Memory Tasks

Because the mathematical and visuo-spatial operations tasks are being used for the first time, the specificity of the locus of the processing load is unknown. Furthermore, because these tasks (in particular, visuo-spatial operations) involve sequences of operations that are unfamiliar to participants, they require conscious processing which may lead to variability in response times across regions. For these reasons, performance on these tasks will be analyzed using what are expected to be ‘critical regions’ of processing load and, in addition, aggregate scores calculated across all regions of each task.

2.4.5.1 Mathematical Operations

Performance on the mathematical operations task was scored based on the time elapsed between button presses (a measure of processing time of each part of equation) and the accuracy of responses. Initially, response times for each part of the equation were collected; however, for the purpose of analysis, parts of the equations were grouped into the following regions:

\[
\begin{align*}
9 &+ (8 - 5) + 4 = \_\_ \\
9 &+ 3 + 1 - 3 = \_\_ \\
\text{Region:} &\quad 1 \quad 2 \quad 3 \quad 4
\end{align*}
\]

The main region of interest was region three, as this region differentiated the ‘SO’ and ‘OS’ versions of the equations. That is, region three was expected to involve the greatest processing and storage demands in the SO equations.

2.4.5.2 Visuo-spatial Operations

The dependent variables for the visuo-spatial operations task were response time, measured by the intervals between button pressing (a measure of processing time), and response accuracy. Response times to each of the three screens (or regions) in each operation
were compared across the 'subject-object' and 'object-subject' conditions. The critical region of interest was hypothesized to be region two of the SO version, which was expected to have the greatest processing and storage demands.
CHAPTER 3: RESULTS

3.1 Overview

The vast majority of studies investigating working memory and sentence processing have analyzed their results by dividing subjects into two or three groups (high, low and sometimes medium) according to working memory scores (e.g., King & Just, 1991, Waters & Caplan, 1996b, MacDonald, Just & Carpenter, 1992). In a recent article, Caplan and Waters (1999) pointed out that while previous literature has correlated working memory capacity and performance on higher level tasks, such as reading, “almost no research into the role of working memory in sentence processing has used correlational analyses” (p. 79). The working memory measures designed for this study (mathematical and visual-spatial operations) yield reaction time data, which are continuous and therefore cannot be broken into groups in a non-arbitrary way. For this reason, the data from this study will be analyzed using correlation and step-wise regression. The following describes the manner in which the data were prepared for analysis, the analyses used and the results of this experiment.

3.2 Preparation of the Data for Analysis

Reaction time data were trimmed using a procedure adapted from Tyler (1992). First, descriptive statistics (means and standard deviations) were calculated for each condition. Clear outliers, defined as any values lying more than four standard deviations above or below the mean, were changed to missing values, and the mean and standard deviations for each condition were recalculated. Any values lying three standard deviations above or below the mean were then replaced with those maximum or minimum values. The mean and standard deviations per condition were recalculated a final time, and missing values were then
replaced with the new mean value. This procedure resulted in the trimming and replacement of less than five percent of the data in each condition.

For the Waters and Caplan reading span test, reaction times for all responses (correct and incorrect) were trimmed by first identifying the clear outliers at more than three standard deviations from the mean and changing them to missing values. The mean and standard deviation were then recalculated, and any values lying two standard deviations beyond the mean were replaced with that maximum value (resulting in replacing a total of 4.8% of the data). The missing values were then replaced with the new mean.

A composite score for performance on the Waters-Caplan reading span was obtained following the procedure outlined in Waters and Caplan (1996a). Slight alterations were made to the procedure, due to differences in test administration. Reaction times for both correct and incorrect judgements, percentage correct on sensibility judgements and reading span (recall scores) were converted to z-scores. These three z-scores were then averaged, to obtain a final metric in which reaction time, accuracy of judgements and accuracy of recall were all equally weighted. The data for one subject were excluded from this calculation because reaction time data were missing due to a technical error during testing.

3.3 Syntactic Complexity and Syntactic Ambiguity

Participants were considered to have attempted to interpret each sentence irrespective of whether they were successful in answering the comprehension question. That is, since participants were aware that a question would be asked following each sentence, it is likely that they were processing for comprehension. As long as participants were attempting to interpret the sentences, no difference was expected between reaction times for sentences that were correctly interpreted, and those that were not, as measured by the comprehension
question that followed each sentence. To test this hypothesis, for all sentence processing variables (i.e., complexity and ambiguity), comparisons were carried out first on only those sentences for which the subject correctly responded to the comprehension question, then on all sentences, regardless of accuracy on the comprehension measure. No differences were found between the two analyses, and so the analyses reported include response times to all sentences.

3.3.1 Syntactic Complexity: Object-Subject vs. Subject-Object Sentences

The syntactic complexity variable compared participants' performance on processing and comprehending object-subject (OS) and subject-object (SO) sentences. Group response time means for each region of these two sentence types are presented in Table 1.

Table 1

<table>
<thead>
<tr>
<th>Syntactic Complexity Variable: Group (N = 30) Response Time Means and Standard Deviations by Region (in msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>V 1</td>
</tr>
<tr>
<td>V 2</td>
</tr>
<tr>
<td>NP 1</td>
</tr>
<tr>
<td>NP 2</td>
</tr>
<tr>
<td>Final phrase</td>
</tr>
</tbody>
</table>

Note. V 1 = first verb phrase; V 2 = second verb phrase; NP 1 = first noun phrase; NP 2 = second noun phrase.

It was hypothesized that participants would have more difficulty processing the SO constructions than the OS constructions. A 2 × 5 repeated measures ANOVA was carried out to test this hypothesis, comparing sentence type (object-subject vs. subject-object) by region (verb 1, verb 2, NP 1, NP 2, final phrase). The results of this ANOVA are presented in Figures 1 and 2. Figure 1 shows mean response times by sentence type, whereas Figure 2
depicts the differences in response times across sentence types. The main effects of sentence type, $F(1, 29) = 64.19, \text{MSE} = 1,118,443, p = .0001$ and region, $F(4, 29) = 38.35, \text{MSE} = 890,729, p = .0001$ were significant. There was also a significant interaction between sentence type and region, $F(4, 116) = 33.28, \text{MSE} = 401,542, p = .0001$. Planned pairwise comparisons were conducted and interpreted based on Holm’s modified Bonferroni procedure (Shaffer, 1995). They revealed that Verb 1 and Verb 2 of the SO sentences took significantly longer to process than the analogous OS regions, $t(1, 29) = 6.18, p = .0001$ and $t(1, 29) = 8.12, p = .0001$, respectively. Thus, the SO sentences did take longer for participants to process, particularly in the regions which were hypothesized to carry the greatest processing and storage load.

![Reaction times by region for Object-subject and Subject-object sentences](image)

**Figure 1.** Mean reaction times (and standard errors) for each phrase of the complex object-subject and subject-object sentences. V1 = first verb; V2 = second verb; NP1 = first noun phrase; NP2 = second verb phrase; Final = final phrase of the sentence.
Figure 2. Mean reaction time differences (and standard errors) between subject-object and object-subject sentences for each phrase of the sentence. V1 = first verb; V2 = second verb; NP1 = first noun phrase; NP2 = second verb phrase; Final = final phrase of the sentence.

Reaction times for Verb 1 and Verb 2 were aggregated into one region for both the SO and OS constructions for further analyses using correlation and regression with the working memory measures (see sections 3.6 and 3.7). This decision is supported by the data, in which these regions showed comparable reaction times. Furthermore, it is theoretically motivated in that these regions are expected to carry a greater processing load; in the SO constructions, when the first verb is encountered, the subject must assign the thematic role ‘agent’ to the subject of the embedded clause, and ‘patient’ to the first noun phrase. The subject is also made aware that the first noun phrase must be further retained in memory until the main verb is encountered. When the main verb is encountered, this final thematic role dependency can be resolved and the role agent can be assigned to the first noun phrase. Thus, the greatest processing load is expected for both the first and the second verbs in the SO constructions.
3.3.2 Syntactic Ambiguity

The group response time means and standard deviations by region for the main verb and reduced relative ambiguous and unambiguous sentences are summarized in Table 2.

Table 2

<table>
<thead>
<tr>
<th>Region</th>
<th>Main verb</th>
<th>Reduced relative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ambiguous</td>
<td>Unambiguous</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>1</td>
<td>565</td>
<td>152</td>
</tr>
<tr>
<td>2</td>
<td>567</td>
<td>146</td>
</tr>
<tr>
<td>3</td>
<td>842</td>
<td>341</td>
</tr>
</tbody>
</table>

Note. Region 1 starts at the verb and introduces the ambiguity in ambiguous sentences (e.g., *The soldiers warned about the dangers*). Region 2 is the continuation of the phrase in main verb sentences (*before the midnight*), Region 2 is the main verb phrase in reduced relative sentences (*conducted the midnight*), Region 3 corresponds to the sentence final word (*raid*).

3.3.2.1 Main Verb Sentences

It was hypothesized that temporarily ambiguous main verb sentences (MVA) would take longer for all participants to process than unambiguous main verb sentences (MVU). A 2 x 3 repeated measures ANOVA for the ambiguous and unambiguous main verb sentences revealed a marginally significant main effect of sentence type, \( F (1, 29) = 3.37, \text{MSE} = 27,494, p = .08 \). Interactions between sentence type and region were not significant, \( F (2, 58) = 1.02, \text{MSE} = 5267, p = .37 \). The results from this analysis appear in Figures 3 (mean response times by sentence type) and 4 (response time differences).
Figure 3. Reaction times (and standard errors) for each region of the temporarily ambiguous and unambiguous main verb sentences. Region 1 corresponds to the introduction of the main verb clause (for both ambiguous and unambiguous versions); Region 2 is the continuation of the phrase (and continuation of the ambiguity); Region 3 is the final word of the sentence and is the point of disambiguation for the ambiguous sentences. MVU = main verb unambiguous; MVA = main verb ambiguous.
Figure 4. Mean differences in reaction time (and standard errors) for each region of the main verb ambiguous and main verb unambiguous sentences.

3.3.2.2 Reduced Relative Sentences

It was expected that temporarily ambiguous reduced relative sentences would be more difficult for participants to process than their unambiguous counterparts, resulting in longer reaction times. A $2 \times 3$ repeated measures ANOVA for the temporarily ambiguous and unambiguous reduced relative sentences revealed a main effect of sentence type, $F(1, 29) = 11.95$, MSE = 452,215, $p = .002$ and a main effect of region, $F(2, 29) = 29.84$, MSE = 1,902,719, $p = .0001$. The interaction between sentence type and region was also significant, $F(2, 58) = 8.74$, MSE = 336,991, $p = .0005$. These results are depicted in Figures 5 (mean response times by sentence type) and 6 (response time differences). Follow-up multiple comparisons were conducted and interpreted using Holm’s modified Bonferroni. There was a statistically significant difference between Region 2 of the temporarily ambiguous and unambiguous sentences, $t(1, 29) = 6.46$, $p = .0001$, which is the point of disambiguation for
the ambiguous construction. There were no significant differences between ambiguous and unambiguous sentences at Region 3, the final point of disambiguation.

Figure 5. Reaction times for each region of the temporarily ambiguous and unambiguous reduced relative sentences. Region 1 corresponds to the relative clause; Region 2 corresponds to the main verb clause, which is the first point of disambiguation; Region 3 is the sentence final word, which is the final point of disambiguation. RRU = reduced relative unambiguous; RRA = reduced relative ambiguous.

For use in further analyses (see section 3.6), response times in the second and third regions were aggregated into one region for the main verb and reduced relative constructions. For the main verb ambiguous sentences, the second region was the first possible point of disambiguation, while the third region was the final point of disambiguation. In the case of reduced relative ambiguous sentences, the second region was the point of disambiguation, and some of this effect was still apparent in the third region of the sentence.
3.4 New Working Memory Tasks

Group means of performance on the mathematical operations and the visuo-spatial operations tasks are reported in Table 3. For both of these tasks, it was hypothesized that the SO condition, designed to match the processing and storage requirements of subject-object sentences, would be more difficult to process than the OS condition, which mirrored the demands of object-subject sentences. This would result in longer reaction times overall, and particularly in specific regions of the tasks.
Table 3

*New Working Memory Tasks: Group Response Time Means and Standard Deviations by Region for the Mathematical Operations and Visuo-spatial Operations Tasks (in msec)*

<table>
<thead>
<tr>
<th>Region</th>
<th>Mathematical Operations</th>
<th>Visuo-spatial Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OS M</td>
<td>SD</td>
</tr>
<tr>
<td>1</td>
<td>654</td>
<td>186</td>
</tr>
<tr>
<td>2</td>
<td>1536</td>
<td>513</td>
</tr>
<tr>
<td>3</td>
<td>1078</td>
<td>231</td>
</tr>
<tr>
<td>4</td>
<td>1692</td>
<td>418</td>
</tr>
</tbody>
</table>

Note. There is no Region 4 for the visuo-spatial operations task. OS and SO refer to the sentence type that the working memory task was designed to match (object-subject and subject-object, respectively).

3.4.1 Mathematical Operations

To examine the effects of complexity on the mathematical operations a $2 \times 4$ repeated measures ANOVA compared operation type (OS,SO) and region (1,2,3,4). The results from this analysis are summarized in Figures 7 and 8. The main effect of operation type was not significant, $F (1,29) = .10$, MSE = 2077, $p = .76$; however, there was a significant main effect of region, $F (3,29) = 152.82$, MSE = 11,282,331, $p = .0001$. The interaction between operation type and region was also significant, $F (3,87) = 99.58$, MSE = 7,084,881, $p = .0001$. 
Figure 7. Reaction times (and standard errors) by region for the complex (SO) and simple (OS) conditions of the mathematical operations working memory measure.

Figure 8. Differences in reaction times (and standard errors) for the SO minus the OS conditions of the mathematical operations working memory task.
Follow-up multiple comparisons were carried out to determine the source of the interaction. Region 3 of the SO condition was the main point of interest (refer to Hypothesis 3); it was expected that this region would require more processing time than the analogous region of the OS condition. A t-test comparing these regions confirmed that there was a statistically significant difference, $t(1, 29) = 10.68, p = .0001$. However, significant differences at other compared regions were in the opposite direction; for Region 2, the OS condition took significantly longer to process than the SO condition, $t(1, 29) = -9.13, p = .0001$. The same was true at Region 4, $t(1, 29) = -5.21, p = .0001$.

3.4.2 Visuo-spatial Operations

Effects of complexity for the visuo-spatial operations task were examined using a $2 \times 3$ repeated measures ANOVA which compared operation type (OS, SO) and region (1, 2, 3). The frames used for the visuo-spatial task were the same for both operation types, only the order of presentation changed (refer to Appendix F for examples). Region 1 was one frame presented on the computer screen. This frame appeared on the first screen for the SO condition of the operation and on the second screen for the OS condition. Region 2 involved two frames on the screen at once, and was presented on the second screen for the SO condition, and on the first screen for the OS condition. Region 3 was a decision screen with two frames, and appeared last in both operation types. For post-hoc analysis, the regions were compared across operation types (e.g., Region 1 for OS and SO). Results from the ANOVA are graphically summarized in Figures 9 and 10.
Reaction times by region for Visuo-spatial Operations (OS and SO conditions)

Region

Figure 9. Reaction times (and standard errors) for each region of the simple (OS) and complex (SO) conditions of the visuo-spatial operations working memory measure.

Visuo-spatial Operations:
Subject-object minus Object-subject

Mean reaction time difference (msec)

Region

Figure 10. Mean differences in reaction times (and standard errors) for each region of the SO minus the OS conditions of the visuo-spatial operations working memory task.
There were significant main effects of operation type and of region, $F(1, 29) = 24.06$, $MSE = 19,120,297, p = .0001$ and $F(2, 29) = 70.65, MSE = 44,700,838, p = .0001$, respectively. There was also a significant interaction of operation type and region, $F(2, 58) = 19.36, MSE = 10,264,989, p = .0001$. Follow-up comparisons were carried out to determine the basis of the interaction. Of particular interest was the comparison at Region 2; it was hypothesized that this region in the SO condition of the operation would carry the heaviest processing load, and would therefore take significantly longer to process than Region 2 in the OS condition. In fact, this difference was not significant, $t(1, 29) = 0.80, p = .43$. The difference underlying the significant interaction of operation type and frame was found in the comparison of Region 1 of each condition; Region 1 of the SO condition took significantly longer for subjects to process than the same region in the OS condition, $t(1, 29) = 5.48, p = .0001$. Given this finding, and the fact that participants were aware that Region 1 in the SO condition needed to be stored while computing Region 2, it seemed reasonable that both Region 1 and Region 2 in the SO condition of the visuo-spatial operations contributed to the greater difficulty of this operation type versus the OS condition. For this reason, reaction times for Regions 1 and 2 of the SO condition were aggregated for subsequent analyses (see sections 3.5 and 3.6).

3.5 Correlations among the Working Memory Tasks

Correlations were calculated to determine the extent to which the different working memory measures were related. As mentioned earlier, analyses were conducted by region and as aggregates.
Participants’ raw scores on the Daneman and Carpenter reading span task, the Waters and Caplan reading span task, as well as composite scores for the Waters and Caplan reading span task are listed in Appendix G. Participants’ mean response time scores (and standard deviations) for both the critical regions and the aggregates of the mathematical operations and visuo-spatial operations working memory tasks are listed in Appendix H.

3.5.1 Critical Regions

The results of the correlation analysis are presented in Table 4. It was hypothesized that correlations among the working memory tasks would vary in strength, thus reflecting the extent to which the tasks relied on overlapping sequences of processing demands.

Surprisingly, few of the correlations between the working memory measures were significant. The strongest correlation was between the Waters and Caplan reading span task and the composite score ($r = .511, p = .004$), which is not unexpected given that the recall score in the reading span was one of the components of the composite. There was a significant correlation between scores on the Waters and Caplan composite measure and the visuo-spatial operations task ($r = .459, p = .01$). Finally, a marginally significant correlation was found between performance on the mathematical operations task and the visuo-spatial operations task ($r = .389, p = .04$). All other correlations fell below the level of significance.
Table 4

**Correlation Matrix of Performance on Working Memory Measures**

<table>
<thead>
<tr>
<th></th>
<th>DC span</th>
<th>WC span</th>
<th>WC comp.</th>
<th>Math</th>
<th>Visuo-spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC span</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC span</td>
<td>.200</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC comp.</td>
<td>-.046</td>
<td>.511**</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td>-.203</td>
<td>-.070</td>
<td>.145</td>
<td>.459*</td>
<td>.389*</td>
</tr>
<tr>
<td>Visuo-spatial</td>
<td>-.355</td>
<td>.188</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. DC span = Daneman & Carpenter's (1980) reading span; WC span = Waters & Caplan's (1996a) reading span; WC comp. = Waters & Caplan's (1996a) composite scores; Math = Mathematical operations, SO condition, Region 3; Visuo-spatial = Visuo-spatial operations, SO condition, combining Regions 1 and 2.*

* p < .05  
** p < .01  

3.5.2 Aggregate Scores

In order to further investigate the relationship between the working memory measures, reaction time scores on the mathematical operations and visuo-spatial operations tasks were aggregated across all regions in each condition. The aggregates were then used in calculating the correlations between these working memory tasks and the reading span working memory measures. The results of these correlations are presented in Table 5.

When aggregate scores for reaction times on the mathematical and visuo-spatial operations tasks are used, these two tasks show a stronger correlation with one another (r = .500, p = .005) as compared with the correlation between the tasks when specific regions were entered into the correlation (r = .389, p = .04). Also note that when aggregate scores are used for the new working memory tasks, these tasks no longer significantly correlate with the
Waters and Caplan composite score for working memory (r = .332, p = .08 for math and r = .356, p = .06 for visuo-spatial operations).

Table 5

Correlations of Performance on all Working Memory Measures Using Aggregate Scores for Mathematical Operations and Visuo-Spatial Operations

<table>
<thead>
<tr>
<th></th>
<th>DC span</th>
<th>WC span</th>
<th>WC comp.</th>
<th>Math</th>
<th>Visuo-spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC span</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC span</td>
<td>.200</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC comp.</td>
<td>-.046</td>
<td>.511**</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td>-.247</td>
<td>-.023</td>
<td>.332</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Visuo-spatial</td>
<td>-.356</td>
<td>.167</td>
<td>.356</td>
<td>.500**</td>
<td>--</td>
</tr>
</tbody>
</table>

**Note.** Reaction times were averaged across all regions of the SO conditions of the mathematical and visuo-spatial operations working memory tasks to obtain an aggregate score for each measure. DC span = Daneman & Carpenter's (1980) reading span; WC span = Waters & Caplan's (1996a) reading span; WC comp. = Waters & Caplan's (1996a) composite scores; Math = Mathematical operations aggregate scores; Visuo-spatial = Visuo-spatial operations aggregate scores.

**p < .01

3.6 Correlations between Working Memory Measures and Specific Sentence Regions

3.6.1 Syntactic Complexity

3.6.1.1 Correlations with Critical Regions of Working Memory Measures

In order to determine the extent to which working memory capacity predicted ability to process complex sentences, correlations were calculated between the various working memory measures and the region of the subject-object and object-subject sentences which was expected to carry the greatest processing load (an aggregate of Verb 1 and Verb 2). These analyses were first carried out using hypothesized ‘critical regions’ of the new
working memory measures, and then with aggregate scores for these measures (see section 3.6.1.2 below). Results from the first set of correlations are summarized in Table 6.

Table 6

Correlations between Working Memory Measures (Critical Regions of Mathematical and Visuo-spatial Operations) and Critical Regions of the Object-subject and Subject-object Sentences

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>DC span</th>
<th>WC span</th>
<th>WC comp.</th>
<th>Math</th>
<th>Visuo-spatial 1</th>
<th>Visuo-spatial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>.170</td>
<td>.341</td>
<td>.564**</td>
<td>.418*</td>
<td>.225</td>
<td>.485**</td>
</tr>
<tr>
<td>SO</td>
<td>.149</td>
<td>.208</td>
<td>.471**</td>
<td>.439*</td>
<td>.319</td>
<td>.573**</td>
</tr>
</tbody>
</table>

Note. DC span = Daneman & Carpenter reading span; WC span = Waters & Caplan reading span; WC comp. = Waters & Caplan composite scores; Math = Mathematical operations, SO condition, Region 3; Visuo-spatial 1 = Visuo-spatial operations, SO condition, Region 2; Visuo-spatial 2 = Visuo-spatial operations, SO condition, combining Regions 1 and 2; OS = Object-subject sentence, aggregate of Verb 1 and 2; SO = Subject-object sentence, aggregate of Verb 1 and 2.

* p < .05

** p < .01

3.6.1.1.1 Mathematical Operations

The region of the mathematical operations task requiring the greatest amount of processing (Region 3 of the SO condition) correlated significantly with the region of heaviest processing of the subject-object and object-subject sentences (r = .439, p = .04 and r = .418, p = .02, respectively). These findings support hypothesis 5(b), that specific regions of the mathematical operations task correlate strongly with the analogous regions of the sentence processing task.
3.6.1.1.2 Visuo-spatial Operations

No significant correlations were found between the region of the visuo-spatial task involving the greatest processing load (hypothesized to be Region 2 of the SO condition) and the analogous region of the subject-object sentences \((r = .319, p = .09)\) or the object-subject sentences \((r = .225, p = .23)\). Thus, hypothesis 5(c) is not supported.

As discussed earlier (section 3.4.2), combining the first and second regions of the SO condition of the task may provide a more appropriate measure of both storage (Region 1) and processing/integration (Region 2). The correlations were thus recalculated using an aggregate of Regions 1 and 2. For the SO condition, correlations with Verb 1 and Verb 2 of the subject-object and object-subject sentences were significant \((r = .573, p = .0007\) and \(r = .485, p = .006\), respectively).

3.6.1.1.3 Reading Span Tasks

The Daneman and Carpenter reading span task did not correlate significantly with specific regions of the complex sentences which carried the greatest processing load (aggregate of Verb 1 and 2), contrary to what was predicted in the hypotheses. The same pattern of results was found with the Waters and Caplan reading span task when only recall scores were taken into account. However, when the composite score (including reaction time to the grammatical judgement, accuracy of the grammatical judgement and accuracy of recall) was entered into the correlation, it did correlate significantly with reading times of Verb 1 and Verb 2 in the subject-object \((r = .471, p = .009)\) and object-subject sentences \((r = .564, p = .001)\). Thus, partial support is provided for the hypothesis that reading span would correlate with language processing skills, as measured by specific regions in the complex sentences.
3.6.1.2 Correlations with Aggregated Working Memory Measures

The fact that the Waters and Caplan composite scores showed significant correlations with sentence processing, as did the mathematical operations and visuo-spatial operations tasks, suggests that storage and processing demands over the entire working memory task merit further examination. It is possible that the on-line measure used was not sufficiently sensitive to the moment-to-moment demands of each working memory task, or that a greater degree of conscious processing was involved leading to subject variability in response times across regions. For this reason, reaction times in all regions on the mathematical operations and visuo-spatial operations tasks were aggregated to obtain a measure of the storage and processing demands across each task.

Correlations between the aggregated working memory scores and the critical regions of interest in sentence processing were recalculated. This critical region of the subject-object sentences was hypothesized to correlate strongly with the new working memory measures, but show weaker correlations with more traditional measures of working memory. The results of this analysis are presented in Table 7.

Overall, the correlations are stronger using the aggregate scores for mathematical and visuo-spatial operations (compare with results in Table 6). The strongest correlations are between the aggregate scores for mathematical operations and sentence processing (r = .607, p = .0003 with the object-subject sentences, r = .531, p = .002 with the subject-object sentences). The aggregate scores for the visuo-spatial operations task are also significantly correlated with processing of the complex sentences (r = .437, p = .015 with the object-subject sentences, r = .503, p = .004 with the subject-object sentences).
Table 7

*Correlations between Mathematical and Visuo-spatial Operations Aggregate Scores and Critical Regions of the Object-subject and Subject-object Sentences*

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Math</th>
<th>Visuo-spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>.607**</td>
<td>.437*</td>
</tr>
<tr>
<td>SO</td>
<td>.531**</td>
<td>.503**</td>
</tr>
</tbody>
</table>

Note. The math and visuo-spatial measures are aggregates of reaction time scores for the SO conditions across all regions in each task. Math = Mathematical operations; Visuo-spatial = Visuo-spatial operations; OS = Object-subject sentence, aggregate of Verb 1 and 2; SO = Subject-object sentence, aggregate of Verb 1 and 2.

* p < .05

** p < .01

3.6.2 Syntactic Ambiguity

3.6.2.1 Correlations with Critical Regions of Working Memory Measures

Similar correlational analyses were carried out to determine the extent to which the different working memory measures covaried with performance on processing of syntactically ambiguous sentences. At the outset of the experiment, it was predicted that both the Daneman and Carpenter and the Waters and Caplan reading span tasks would correlate significantly with processing at critical regions of these sentence types. On the other hand, the mathematical operations and visuo-spatial operations tasks were not designed to match the demands of syntactically ambiguous sentences, and therefore, they were not expected to significantly correlate with performance on these sentence types. Results of the correlational analyses are presented in Table 8.
Table 8

*Correlations between the Working Memory Measures and Regions 2 and 3 of the Main Verb and Reduced Relative Ambiguous and Unambiguous Sentences*

<table>
<thead>
<tr>
<th>Sentence</th>
<th>DC span</th>
<th>WC span</th>
<th>WC comp.</th>
<th>Math</th>
<th>Visuo-spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVU</td>
<td>.169</td>
<td>.349</td>
<td>.448*</td>
<td>.221</td>
<td>.405*</td>
</tr>
<tr>
<td>MVA</td>
<td>.198</td>
<td>.340</td>
<td>.335</td>
<td>.070</td>
<td>.267</td>
</tr>
<tr>
<td>RRU</td>
<td>.183</td>
<td>.413*</td>
<td>.560**</td>
<td>.269</td>
<td>.459**</td>
</tr>
<tr>
<td>RRA</td>
<td>.075</td>
<td>.298</td>
<td>.292</td>
<td>.190</td>
<td>.358**</td>
</tr>
</tbody>
</table>

Note. The mathematical operations measure in this analysis is Region 3 of the SO condition. The visuo-spatial measure used in this analysis combines Region 1 and Region 2 of the SO condition. MVU = main verb unambiguous; MVA = main verb ambiguous; RRU = reduced relative ambiguous; RRA = reduced relative ambiguous.

* p < .05
** p < .01

Overall, few of the correlations reached significance. None of the correlations between the Daneman and Carpenter reading span task and sentences in the syntactic ambiguity variable were significant, contrary to expectation. The Waters and Caplan reading span task did correlate significantly with the reduced relative unambiguous sentences (r = .413, p = .02); however, it did not correlate significantly with any of the other sentence types. The Waters and Caplan composite scores for working memory correlated significantly with both main verb and reduced relative sentences, but only for performance on the unambiguous conditions of these sentence types (r = .560, p = .001 and r = .448, p = .014, respectively).

None of the correlations between mathematical operations and sentences for the syntactic ambiguity variable were significant, in accordance with the hypotheses. There were significant correlations between performance on the SO condition of the visuo-spatial task and both the unambiguous and ambiguous conditions of the reduced relative sentences (r =
.459, p = .009 and r = .358, p = .05) and for the unambiguous main verb sentences (r = .405, p = .03). These findings are contrary to what was predicted in the hypotheses.

3.6.2.2 Correlations with Aggregated Working Memory Measures

Correlations between working memory measures and critical regions of the syntactically ambiguous sentences were recalculated using aggregate scores for the new working memory tasks. The results of this analysis are presented in Table 9.

Table 9

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Math</th>
<th>Visuo-spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVU</td>
<td>.347</td>
<td>.324</td>
</tr>
<tr>
<td>MVA</td>
<td>.192</td>
<td>.192</td>
</tr>
<tr>
<td>RRU</td>
<td>.408*</td>
<td>.374*</td>
</tr>
<tr>
<td>RRA</td>
<td>.272</td>
<td>.315</td>
</tr>
</tbody>
</table>

* p < .05

** p < .01

There were no significant correlations between the mathematical and visuo-spatial operations tasks and main verb sentences. With respect to the reduced relative sentences, there were no significant correlations between these new working memory measures and the temporarily ambiguous condition of the reduced relative sentences. Only one of the correlations involving mathematical operations and visuo-spatial operations was significant. Mathematical operations (but not visuo-spatial operations) correlated significantly with the unambiguous reduced relative sentences (r = .408, p = .02). This pattern of results contrasts with the earlier analysis correlating specific regions of these working memory tasks with
these sentence types, in which visuo-spatial operations, but not mathematical operations, correlated significantly with both conditions of the reduced relative sentences (compare with results summarized in Table 8).

3.7 Regression of the Working Memory Measures on Performance on Complex Sentences

A stepwise regression analysis was carried out to further examine the relative predictive ability of the working memory measures on performance on the complex subject-object sentences. The dependent variable for this analysis was reaction time at the critical region of the subject-object sentences (an aggregate of Verb 1 and Verb 2). The independent variables entered into the stepwise regression were aggregate scores for the mathematical and visuo-spatial operations tasks, the Daneman and Carpenter reading span, the Waters and Caplan reading span and the composite scores from Waters and Caplan’s working memory measure.

At the outset of the experiment, it was hypothesized that the mathematical and visuo-spatial operations tasks would account for a greater amount of the variance in performance on these sentence types than would the reading span tasks. The results of this analysis are presented in Table 10.
Table 10

**Summary of Stepwise Regression Analysis for Variables Predicting Ability to Process Complex Subject-object Sentences**

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \Delta R^2 )</th>
<th>B</th>
<th>SE B</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematical operations</td>
<td>.27*</td>
<td>.620</td>
<td>.196</td>
<td>.519*</td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematical operations</td>
<td>.10*</td>
<td>.487</td>
<td>.197</td>
<td>.408*</td>
</tr>
<tr>
<td>WC composite scores</td>
<td>172.590</td>
<td>85.111</td>
<td></td>
<td>.335*</td>
</tr>
</tbody>
</table>

Note. The mathematical operations variable is based on aggregate scores across all regions of this task. WC = Waters & Caplan composite scores.

*p < .05

This analysis revealed that mathematical operations accounted for 27% of the variance in the performance on complex subject-object sentences. Thus, mathematical operations was the strongest predictor of variables entered into the regression. The Waters and Caplan composite scores accounted for a further 10% of the variance in the sentence processing scores. None of the other variables (Daneman and Carpenter reading span, Waters and Caplan reading span, visuo-spatial operations) were able to provide further predictive ability, and therefore were not entered into the model.

3.8 Question Accuracy and Decision Times

3.8.1 Accuracy of Performance on the Comprehension Questions

The overall percentage of comprehension accuracy for each sentence type as well as on each working memory measure is summarized in Table 11.
Table 11

*Overall Accuracy on Sentence Comprehension Questions and Working Memory Tasks*

<table>
<thead>
<tr>
<th>Sentence Comprehension</th>
<th>Working Memory&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence Type</td>
<td>Percent Correct</td>
</tr>
<tr>
<td>Object-Subject</td>
<td>90</td>
</tr>
<tr>
<td>Subject-Object</td>
<td>78</td>
</tr>
<tr>
<td>MVA</td>
<td>87</td>
</tr>
<tr>
<td>MVU</td>
<td>84</td>
</tr>
<tr>
<td>RRA</td>
<td>46</td>
</tr>
<tr>
<td>RRU</td>
<td>72</td>
</tr>
</tbody>
</table>

<sup>a</sup> Percent correct refers to accuracy on the processing component (and not recall) on the working memory tasks. The Daneman-Carpenter reading span task did not incorporate a direct measure of processing accuracy.

For the syntactic complexity variable, participants made more errors in comprehension on the subject-object (accuracy = 78%) than on object-subject constructions (accuracy = 90%). This provides further evidence that the former was a more demanding construction for participants. With respect to the syntactic ambiguity variable, participants made fewer errors on the main verb constructions than the reduced relative constructions. Counterintuitively, subjects were slightly more accurate in their interpretation of main verb ambiguous constructions (accuracy = 87%) than main verb unambiguous constructions (accuracy = 84%). Comprehension accuracy of the reduced relative clause sentences was much lower overall; accuracy on the reduced relative ambiguous sentences was at chance levels (46%) and accuracy on the unambiguous reduced relative sentences was 72%. A one-way repeated measures analysis of variance comparing comprehension accuracy across sentence type showed a significant main effect of sentence type, $F(5, 145) = 48.98$, $MSE = 7934$, $p = .0001$, confirming that certain sentence types were more difficult to process than others.
Turning to the working memory measures, overall there was a high level of accuracy on each task (the lowest accuracy level was 87% for the SO condition of the visuo-spatial operations task). The mathematical and visuo-spatial operations tasks showed slightly lower levels of accuracy for the more complex SO conditions than the OS conditions. A one-way repeated measures ANOVA revealed a significant main effect of working memory measure, $F (4, 112) = 5.08, \text{MSE} = 336, p = .0008$, confirming that participants were more accurate in some working memory tasks than others. Follow-up pairwise comparisons determined that there were no significant differences in accuracy across OS and SO conditions for both the mathematical and visuo-spatial operations tasks. This is likely due to the fact that there were no differences in complexity at the final part of each operation. Accuracy on the Waters and Caplan reading span task was significantly higher than accuracy on the SO conditions for mathematical and visuo-spatial operations, and the OS condition of mathematical operations ($p = .05$). There was a marginally significant difference between the Waters and Caplan reading span task and the OS condition of visuo-spatial operations ($p = .06$). Thus, participants were more accurate on the Waters and Caplan reading span task than on the new working memory tasks.

Correlations were calculated between accuracy on the working memory tasks and accuracy on sentence processing. Significant correlations were observed between accuracy on the SO condition of the visuo-spatial operations task and accuracy on both the OS and SO sentences ($p = .005$ and $p = .003$, respectively). No other correlations reached significance.

3.8.2 Response Times to the Comprehension Questions

An analysis of variance was carried out to examine decision time to the comprehension question by sentence type. There was a significant main effect of sentence type, such that
participants took longer to answer comprehension questions for certain sentences, \( F(5, 145) = 10.20, \text{MSE} = 6,753,795, p = .0001 \). Follow-up multiple comparisons determined that participants took significantly longer to answer questions for both the temporarily ambiguous and the unambiguous reduced relative sentences as compared with their main verb counterparts, \( t(1, 29) = 3.38, p = .002 \) and \( t(1, 29) = 4.42, p = .0001 \), respectively. Participants also took significantly longer to respond to questions following subject-object versus object subject sentences, \( t(1, 29) = 3.09, p = .004 \).

3.9 Summary of Results

In summary, main effects of complexity and of ambiguity (in reduced relative sentences, but not main verb sentences) were observed. The new working memory tasks showed main effects of operation type (OS versus SO). Significant correlations were observed between response times in critical regions of the complex subject-object sentences and mathematical operations, visuo-spatial operations and Waters and Caplan’s composite reading span scores.
CHAPTER 4: DISCUSSION

4.1 Introduction

This study examined the role of working memory capacity in modulating ability to comprehend syntactically complex and syntactically ambiguous sentences. Previous studies in this area have yielded conflicting results (e.g., King & Just, 1991; MacDonald et al., 1992; Caplan & Waters, 1999; Caplan & Waters, 1996b). It has been suggested that the conflicting results may be due to the use of off-line measures of sentence processing and different measures of working memory capacity across studies. To overcome these problems, this study employed an on-line, self-paced reading task to measure sentence processing. To ensure continuity with past studies, the most widely-used reading span task, developed by Daneman and Carpenter (1980), was used, along with a more recently developed task from Waters and Caplan (1996a). A further goal of the study was to examine the nature of working memory itself, and whether tasks in different domains tap into a single resource or domain-specific resources. More specifically, it sought to determine whether the ability of a working memory test to predict ability in sentence processing was more reliant on the processing characteristics of the working memory task or the domain (e.g., visuo-spatial) of that task. Thus, two new measures of working memory, a mathematical operations and a visuo-spatial operations task, were developed to match the sequence of processing demands in syntactically complex sentences. From a connectionist perspective, it is the patterns of activation of the task, and not the domain of that task, which are paramount (cf. Christiansen & Chater, 1999). Therefore, it was expected that stronger correlations would emerge between performance on the complex sentences and working memory tasks designed to match the processing characteristics of complex sentences (mathematical and visuo-spatial operations)
than between sentence processing and more traditional reading span tasks. These new tasks were also hypothesized to show greater relative contribution to the variance in performance on the processing of syntactically complex sentences.

In the discussion that follows, the role of working memory in modulating sentence processing and the importance of domain-specificity in predicting sentence processing ability are addressed. The results presented in chapter three are discussed in relation to the hypotheses outlined in chapter one (section 1.3). Finally, a more general discussion of the implications of the results of this study is presented.

4.2 Overall Performance on Sentence Processing

4.2.1 Syntactically Complex Sentences

It was hypothesized at the outset of the study that understanding syntactically complex (subject-object) sentences would be more difficult for all subjects than the less complex object-subject sentences, as measured by reaction times at critical regions and comprehension errors (see section 1.3.1, hypothesis 1). Results from the study support this research hypothesis. Participants took longer to process the first and second verbs of the subject-object (SO) constructions than the same regions of the object-subject (OS) constructions. They also took significantly longer to respond to comprehension questions following the more complex SO sentences and made more errors when responding to questions following this sentence type. These results replicate previous literature which has also found the first and second verbs of SO sentences to carry a significant processing load (cf. King & Just, 1991; Caplan & Waters, 1999; Small et al., 2000). While researchers generally agree that SO sentences are more difficult to process than OS sentences, they differ in explanations of individual
differences in responding to these sentence types. This will be further discussed in relation to performance on the working memory tasks in section 4.5.1 below.

4.2.2 Syntactically Ambiguous Sentences

It was hypothesized that all participants would have greater difficulty processing ambiguous versus unambiguous sentences, as measured by response times and comprehension questions (see section 1.3.1, hypothesis 2). Based on previous research, it was expected that the second and third regions of the ambiguous sentences (which are the main points of disambiguation) would yield elevated response times in comparison to their unambiguous counterparts (MacDonald et al., 1992; Kemtes & Kemper, 1997). The findings for main verb and reduced relative sentences will be discussed in turn.

4.2.2.1 Main Verb Sentences

With respect to the temporarily ambiguous and unambiguous sentences with main verb interpretation (e.g., *The soldiers warned about the dangers before the midnight raid.* vs. *The soldiers spoke about the dangers before the midnight raid.*), the results did not confirm the prediction. Regions two and three did not take significantly longer to process in the ambiguous versus the unambiguous sentences. Moreover, there was no difference between response times to questions following both sentence types, and participants actually made more errors in responding to unambiguous main verb sentences than the ambiguous sentences. This finding is contrary to results in previous research, using the same sentences (MacDonald et al. 1992; Kemtes & Kemper; 1997). This result lends some support to more recent work which claims that ambiguity resolution may be lexically and not syntactically driven (MacDonald, Pearlmutter & Seidenberg, 1994; Pearlmutter & MacDonald, 1995). These researchers claim that probabilistic knowledge of the characteristics of particular
noun-verb combinations may bias the reader towards a certain interpretation in ambiguous constructions. The stimuli used in this study were created before this explanation of ambiguity resolution was advanced, and did not control for the plausibility of alternative interpretations or the frequency of particular noun-verb combinations with certain sentence structures. Therefore, there may have been lexical ‘cues’ in the stimuli, leading the readers to have no more difficulty with ambiguous than unambiguous sentences. For example, for nearly all of the sentences, the first noun phrase was animate, and typically, when the first noun of a sentence is animate it is the agent of the verb which follows. Thus, readers may have been biased towards a main verb interpretation of the ambiguous sentences, which may have resulted in this sentence type being no more difficult than the unambiguous sentences with main verb interpretation. A second possibility is that there was not enough power to detect a difference between the sentence types; previous studies (MacDonald et al., 1992; Kemtes & Kemper, 1997) have found a much smaller effect size for the main verb sentences than the reduced relative sentences. Because of the small sample size in the present study (thirty participants), there may not have been enough power to detect a relatively small difference between performance on ambiguous versus unambiguous main verb sentences.

4.2.2.2 Reduced Relative Sentences

The ambiguous sentences with reduced relative interpretation were also expected to result in elevated response times at regions two and three and more comprehension errors as compared to the unambiguous versions of these sentences (see section 1.3.1, hypothesis 2). The data from this study support this hypothesis; the critical regions of the ambiguous sentences did take significantly longer for all participants to process. Furthermore, participants required significantly more time to respond to comprehension questions
following the ambiguous sentences with reduced relative interpretation. Comprehension of these sentence types was significantly lower, with participants responding correctly to only 46% of the questions following ambiguous reduced relative sentences, and 72% for their unambiguous counterparts. This general pattern of results follows the findings in previous research (MacDonald et al., 1992; Kemtes and Kemper, 1997). Performance on these sentences will be considered in relation to participants’ performance on the working memory tasks in section 4.5.2.

4.3 Overall Performance on the New Working Memory Tasks

4.3.1 Mathematical Operations

Two versions of the mathematical operations task were designed: one in which the sequence of processing demands matched SO sentences, and one matching OS sentences. It was predicted that the more complex SO condition would be more difficult to process overall, resulting in longer response times and increased errors in processing (see section 1.3.2, hypothesis 3). More specifically, it was expected that region three of the SO condition of the task would carry the greatest processing and storage load, similar to the first and second verb of the SO sentences. Therefore, this region was predicted to be the locus of difficulty in processing when compared to the same region in the OS condition of the task. This hypothesis was supported by the data; participants took significantly longer to process region three of the SO compared with the OS equation. It should be noted, however, that there were significant differences between the two conditions at region two and region four which took longer to process in the OS condition than in the SO condition. These differences are not of theoretical interest, and can be easily explained based on the sequences of processing demands in each task. At region two, participants must complete an operation in
the OS task, whereas for SO, participants are made aware by the open brackets that they must continue to store region one and region two until they encounter the closed brackets. The need to store these regions is made explicit by the brackets, which may encourage participants to use a strategy of speeding up through this region in order to reduce the distance over which storage is required. Moreover, even when the region with the longest response time in each condition (region three for SO and region two for OS) are compared, the SO condition takes significantly longer to process, suggesting that this is indeed a more demanding operation type.

4.3.2 Visuo-spatial Operations

For the visuo-spatial operations measure of working memory, it was hypothesized that the SO condition would be more difficult to process overall than the OS condition of the task (see section 1.3.2, hypothesis 3). In particular, it was predicted that the second region of the SO condition would result in longer response times as compared with the analogous region in the OS condition because this region in the SO condition involves both the storage of material from the first region presented and processing of information in region two. While there was a significant difference between the operation types in the direction predicted, the data did not support the ‘critical region’ hypothesis; there was no significant difference between the SO and OS conditions at Region 2. Further examination revealed that there was a significant interaction between operation type (OS and SO) and region. This interaction was driven by significantly longer processing times at the first region in the SO operation. Participants knew in advance that the first screen in the SO operations needed to be retained until the final screen. This foreknowledge may have prompted participants to spend extra time encoding the first region of these operations. Because the first region of the SO
condition involved a significant storage component, and processing and integration took place at the second region, an aggregate of regions one and two was used in further analyses. The aggregate analyses showed that the SO condition did take significantly longer to process than the OS condition, thereby confirming hypothesis 3. Response accuracy for visuo-spatial operations (and mathematical operations) was not significantly different across conditions. This is likely due to the fact that at the final point in the operation, there was no complexity difference between the OS and SO conditions.

4.4 Correlations among the Working Memory Measures

It was hypothesized that correlations among the working memory measures (Daneman and Carpenter's reading span, Waters and Caplan's reading span and composite working memory score, mathematical operations and visuo-spatial operations) would vary in strength, reflecting the extent to which the tasks relied on similar sequences of processing demands (see section 1.3.2, hypothesis 4). The mathematical and visuo-spatial tasks were expected to show the strongest intercorrelations, since they were both designed to match the processing characteristics of complex sentences. It was further expected that the two traditional measures of working memory, Daneman and Carpenter's reading span and Waters and Caplan's reading span, would correlate since the procedures for the two tasks are essentially the same, and they purport to measure the same thing—i.e., working memory capacity as measured by recall (note that this is not the case for the Waters and Caplan composite score, which incorporates measures of accuracy and response time).

Correlational analyses were carried out using both the 'critical regions' and aggregate scores across all regions of the mathematical and visuo-spatial operations tasks and recall scores for the Daneman-Carpenter and Waters-Caplan reading span tasks (see section 3.5).
Performance on Daneman and Carpenter’s reading span task did not correlate significantly with any of the other working memory measures used in this study. The Waters and Caplan reading span was significantly correlated only with the Waters and Caplan composite score (i.e., an aggregate of the z-scores for recall, response time and accuracy to a sensibility judgement for the reading span task), a finding which is not unexpected given that reading span recall is one of the components of the composite. This lack of correlation between the Daneman-Carpenter and Waters-Caplan reading span tasks was notable even before statistical analyses were conducted; in many cases, participants with scores ranging in the ‘mid’ to ‘low’ reading span on the Daneman and Carpenter often scored in the ‘high’ range on the Waters and Caplan task—the reverse was also true in several cases. Furthermore, participants reported after completing the two tasks that they found one or the other to be much more difficult. This suggests that each of the tests may have tapped into factors or participant characteristics unrelated to storage and processing demands. For example, for some participants, the auditory feedback present in the Daneman and Carpenter reading span task (from reading sentences aloud) may have interfered with recall. For other participants, a preference for auditory feedback may have enhanced performance on the task. It also suggests that recall alone may not be a particularly stable measure of storage and processing demands (cf. Waters & Caplan, 1996a). The composite score on the Waters and Caplan task, which incorporated accuracy and response time to a sensibility judgement, in addition to recall, correlated significantly with visuo-spatial operations, but not with any of the measures of the mathematical operations task. Thus, even when measures of reading span beyond recall were taken into account, it still did not consistently correlate with the other measures of working memory. In summary, correlation results for the reading span tasks do not support
the hypothesis stated at the outset of the study in that the Daneman-Carpenter and the Waters-Caplan reading span tasks did not correlate with one another, nor did they consistently correlate with the working memory measures involving mathematical and visuo-spatial operations developed for this study.

The correlation analyses between mathematical and visuo-spatial operations support the initial hypothesis that these two tasks should correlate strongly because they were both designed to match the processing characteristics of subject-object sentences. Correlations between the critical regions of the mathematical and visuo-spatial operations tasks were significant, but not as strong as correlations between the aggregated regions of the tasks. As pointed out in section 4.3, the region-specific patterns of response time for the visuo-spatial task did not correspond precisely to the pattern predicted (i.e., that Region 2 would have the longest response times). It is likely that the novelty of the visuo-spatial task, in particular, induced participants to process the task on a more conscious level, which led to greater variability in strategies and response times across regions. This might explain the weaker correlations for the critical regions analysis. It appears then that aggregate scores for the new working memory tasks provide a more accurate measure of on-line processing demands.

4.5 Correlations Between Working Memory and Sentence Processing

4.5.1 Syntactic Complexity

4.5.1.1 Reading Span

Performance on the reading span tasks (Daneman and Carpenter, Waters and Caplan) was hypothesized to significantly correlate with response times at the first and second verbs of subject-object sentences, which represent the region with the heaviest processing load (section 1.3.3, hypothesis 5a). This finding would be consistent with previous research using
the same sentence type (King & Just, 1991). It is presumed to reflect the overlap in the set of
skills required for both tasks, i.e., reading and comprehending sentences (cf. MacDonald &
Christiansen, in press). This hypothesis was only partially supported by the data. Neither of
the reading span tasks which only took recall into account correlated significantly with the
sentence processing measure. This finding is consistent with the theory advanced by Caplan
and Waters (1999) that the on-line ‘interpretive’ stage of sentence processing does not
correlate with tasks (such as reading span) that measure verbal, but not sentence processing,
resources. However, the Waters and Caplan composite score did correlate with ability on the
sentence processing task. The latter finding conflicts with Caplan and Waters’ claim that
“working memory capacity, as measured on a task that emphasizes controlled, conscious
manipulation of verbal information, will not correlate with processing efficiency for any
component of the interpretation process” (1999, p. 93). That is, their working memory
composite score reflects “controlled, conscious manipulation” of sentences and words to be
recalled, yet it correlated significantly with on-line comprehension of complex sentences.
With respect to sentence processing, the evidence for Caplan and Waters’ separate sentence
interpretation resource theory is all based on null results – i.e., no difference in
comprehension between low and high span participants (e.g., Caplan & Waters, 1999; Waters
& Caplan, 1996b); these researchers interpret this lack of significant differences as support
for their separate sentence interpretation resource theory. However, it should be noted that
the studies reported on by Caplan and Waters (1999) have used working memory tasks that
measure recall alone. The findings of the present study demonstrate that such measures do
not correlate with sentence processing ability, but that measures incorporating all aspects of
the task (storage and processing, as does the Waters and Caplan composite score) do correlate strongly with ability to process complex sentences on-line.

4.5.1.2 Mathematical Operations

It was hypothesized that performance on the mathematical operations task would significantly correlate with performance on subject-object sentences; specifically, the regions involving the heaviest processing and storage load for each of these tasks (region three for math and the first and second verbs for SO sentences) were expected to correlate (section 1.3.3, hypothesis 5b). The data supported this hypothesis. When only region three of the more complex (SO) condition of the mathematical operations task was compared with performance at the critical region of SO sentences, the correlations were significant. When aggregate scores for all regions of the mathematical operations task were entered into the correlational analysis, it was even stronger. These findings offer support for the connectionist view of working memory capacity, since the mathematical operations measure was designed specifically to match the sequence of processing demands in subject-object sentences. However, these findings are not inconsistent with Just and Carpenter’s (1992) capacity theory of comprehension. These researchers consider mathematics to involve the manipulation of symbolic information, which is not unlike language processing (see also Daneman & Tardif, 1987). Furthermore, they have pointed out that mathematical operations are often mediated verbally, and so are considered to be a verbal measure. Caplan and Waters’ separate sentence interpretation resource theory cannot account for this pattern of results; they draw a distinction between the resources supporting syntactic processing and other verbal processing, and therefore would not predict a correlation between sentence processing and mathematical operations.
4.5.1.3 Visuo-spatial Operations

Performance on the visuo-spatial operations task was hypothesized to correlate with performance at the critical regions of the complex sentences, again because this task was designed to match the sequential processing demands in the complex sentences (see section 1.3.3, hypothesis 5c). It was hypothesized that Region 2 of the SO condition would be the locus of the heaviest processing load in the visuo-spatial task, and as such would produce the strongest correlations with complex sentences. However, the results showed that Region 2 was not the region in which the heaviest processing load occurred (see section 4.3.2). Therefore, the fact that no significant correlations were found between Region 2 of the more complex visuo-spatial operations and the analogous region of the SO sentences was not unexpected. However, when both region one and two were combined and examined in relation to sentence processing, strong correlations emerged. This captures the combined effects of storage/encoding and processing demands across the task. These results support the hypothesized relationship between visuo-spatial processing and processing of syntactic complexity. The results also support the connectionist approach taken in this study; when the sequence of processing demands are the same across tasks, patterns of performance on the tasks will be the same in spite of differences in cognitive domain. Previous research has not found correlations between visuo-spatial span and standardized measures of language processing (e.g., Daneman & Tardif, 1987; Shah & Miyake, 1996). However, these studies have not attempted to match sequential processing characteristics across tasks. Neither the capacity theory (Just & Carpenter, 1992) nor the separate sentence interpretation resource theory (Caplan & Waters, 1999) predict a correlation between visuo-spatial operations (a nonverbal task) and sentence processing. The results of this study pose a distinct problem for
the claim for a division of those resources supporting verbal (or syntactic) versus nonverbal tasks.

4.5.2 Syntactic Ambiguity

4.5.2.1 Reading Span

Performance on both the Daneman and Carpenter and the Waters and Caplan reading span tasks was expected to correlate significantly with performance on the disambiguating regions of the temporarily ambiguous sentences. This finding would build on past research (e.g., MacDonald et al., 1992) which has found an interaction between span and sentence type (ambiguous versus unambiguous). The data revealed a surprising set of results. Daneman and Carpenter's reading span task did not correlate significantly with any of the sentence types (main verb and reduced relative, ambiguous or unambiguous). The Waters and Caplan reading span (recall-only and the composite score) correlated significantly with the reduced relative unambiguous sentences, but not the ambiguous sentences. The Waters and Caplan composite scores additionally correlated with the main verb unambiguous sentences. Thus, the hypothesis that these tasks would predict ability for ambiguity resolution was not supported; in fact, the data show the opposite result.

One possible explanation for these unpredicted results is the instability of these tasks demonstrated in earlier literature. For example, Waters and Caplan (1996a) demonstrated that reading span scores based on recall alone are highly unstable, and their predictive ability can be altered by minor changes to the sentence types that make up the task. Walenski and Swinney (1999) have criticized the widespread use of the Daneman and Carpenter reading span task on a similar basis; the sentences used in the original version of the task (and possibly subsequent versions) were not controlled for either syntactic complexity or number
of propositions. Thus, it is not unexpected that different versions might yield widely different results. Walenski and Swinney (1999) explain that since these factors are not controlled for, it is unclear how complex sentences are distributed throughout the test. This may result in low or mid span participants receiving scores lower than their actual score (if the more complex sentences are concentrated near the beginning) or mid or high span participants receiving higher scores than they should (if more simple sentences appear at the end). The findings of the present study further attest to the instability of traditional recall-only working memory measures.

4.5.2.2 Mathematical Operations

Correlations between the mathematical operations task and the disambiguating regions of syntactically ambiguous sentences were expected to be weaker than with syntactically complex sentences since the working memory task was not designed to match the sequence of processing demands in ambiguous sentences. Only one out of eight correlations between mathematical operations and the ambiguous and unambiguous sentences was significant. This one correlation was weaker than the correlations with complex sentences, which is consistent with research hypothesis 4b (section 1.10.3).

4.5.2.3 Visuo-spatial Operations

Similar to the mathematical operations task, performance on visuo-spatial operations was hypothesized to correlate weakly (or not at all) with performance on syntactically ambiguous sentences (section 1.10.3, hypothesis 4c). The data support this hypothesis in part; four out of eight possible correlations with syntactically ambiguous and unambiguous sentences were significant, and these correlations were generally weaker than those between visuo-spatial operations and complex sentences.
4.6 Regression of the Working Memory Measures on Performance on Complex Sentences

Performance on the complex sentences was hypothesized to be predicted to varying degrees by the performance on different working memory tasks (section 1.3.3, hypothesis 6). It was expected that the working memory tasks designed for this study (mathematical and visuo-spatial operations) would account for a greater amount of the variance in performance on the complex (subject-object) sentences than would performance on the two reading span tasks. The results from a step-wise regression analysis revealed that mathematical operations was the best predictor of performance on subject-object sentences, accounting for 27% of the variance. The next predictor entered into the model was the Waters and Caplan composite score, which accounted for a further 10% of the variance. None of the other working memory measures could account for additional variance in performance on sentence processing. It is not unexpected that the visuo-spatial task was not entered into the regression model; it strongly correlated with the mathematical operations task ($r = .500$), and thus it would not be expected to account for further variance in the sentence processing task. To determine whether this was the case, another step-wise regression was conducted in which the visuo-spatial task was forced into the model as the first independent variable. When forced into the model, visuo-spatial operations accounted for 25% of the variance in performance at the critical region of SO sentences. Thus, the data support the hypothesis that working memory tasks designed to match the sequence of processing demands in complex sentences will be the best predictors of ability to process those sentences.
4.7 General Discussion

The present study was designed to investigate the following three questions:

1. Do working memory tasks in different domains tap into a single resource?

2. Is the ability to comprehend syntactically complex or syntactically ambiguous sentences modulated by working memory capacity?

3. Does the predictive ability of a working memory task rely less on the domain of processing and more on the similarity of its processing characteristics to the criterion task?

The discussion which follows will address the way in which these questions were answered by the pattern of results observed in the study.

With respect to the first question, few correlations were found among the working memory tasks; the two reading span tasks did not correlate with one another, or with the new working memory tasks. However, the new working memory tasks, measuring processing in different cognitive domains, did correlate strongly with one another when performance across regions of the tasks was taken into account. This suggests that the sequence of processing demands (patterns of activation) may play a more important role than domain of processing in determining performance on the tasks.

There is no one clear answer to the second question, which asks whether working memory capacity modulates sentence comprehension. On one hand, reading span tasks using traditional recall-only scores were not predictive of ability on either of the sentence processing variables in the study. Only when additional measures were combined with recall (in the Waters and Caplan composite) did this become a good predictor of ability to comprehend complex sentences. On the other hand, the new tasks designed to match the
sequence of processing demands in complex sentences proved to be the best predictors of performance on the sentence processing task. This pattern of results lends support to the connectionist view of the nature of “resources” required for sentence processing.

The answer to the final question, whether the predictive ability of a working memory task for sentence comprehension depends less on the domain of processing and more on processing characteristics, is a clear ‘yes’. Indeed, at the outset of the study, such a clear-cut answer to this question was not expected, in that stronger correlations were predicted between reading span and sentence processing. The working memory tasks which involved sentence processing (the traditional recall-only reading span tasks) did not correlate significantly with any of the sentence processing measures used in this study (syntactic complexity or syntactic ambiguity). When a composite score for reading span was employed, it did correlate significantly with sentence processing ability; however, these correlations were still not as strong as those produced by working memory measures in other domains. In spite of differences in domain, the best predictors of ability to process complex sentences were the mathematical and visuo-spatial operations tasks, which were created to have similar processing characteristics to the complex sentences. It should be noted that the mathematical and visuo-spatial operations tasks did not correlate consistently with sentences for which the sequence of processing demands was dissimilar (syntactically ambiguous sentences). This provides further support that sequential processing characteristics are more important than domain of processing in determining the predictive ability of "working memory" tasks.

This study provides an example of how, when correctly designed, a task in one domain can correlate with tasks in another domain. This is not the first time math has been found to correlate with language, but to my knowledge, this is the first time a visuo-spatial measure
has been found to correlate with sentence processing. The latter finding presents a challenge to the claim that there is no question of the dissociation between visuo-spatial tasks and language tasks (see Caplan & Waters, 1999). On the contrary, these findings provide strong support for connectionist models of "working memory capacity", in which processing is not modular or domain-specific, but is rather based on sequences of processing demands which create certain patterns of activation. Through back propagation, connectionist models link current and previous input such that the model can learn from experience (cf. Chater & Christiansen, 1999). Greater exposure to language leads to familiarity with patterns in the input. Differing amounts of experience with language has been proposed to underlie individual differences in sentence processing (MacDonald & Christiansen, in press). In fact, connectionist models encoding “experience” in this way have captured working memory capacity effects described in other studies, without making recourse to a separate “pool of resources” (MacDonald & Christiansen, in press). Models of working memory which incorporate a dissociation between verbal and nonverbal processing resources, or further fractionation between syntactic processing resources and verbal and nonverbal resources, cannot account for the pattern of results found in this study.

On a larger scale, the findings of the present study underline one of the core problems with most models of working memory. “Working memory” is a nebulous ill-defined concept, as are analogous terms such as “processing resources” and “capacity”. As Navon (1984) pointed out, there are numerous flaws with the tenets of resource theory. The most notable of these is the circular argumentation by which individual differences in language processing are said to be due to different-sized resource pools, which can be measured by individual differences in other (working memory) tasks, which are also based on language processing.
The connectionist view on "working memory" advanced by MacDonald and Christiansen (in press; Christiansen & Macdonald, 1999a; 1999b) circumvents this circularity; individual differences in sentence processing ability are a function of experience with language and the architecture of the processing system (i.e., back propagation, hidden unit layer, strength of connections). They argue that the distinction between language tasks and working memory tasks is arbitrary; performance on both will be similar to the extent that each imposes comparable sequences of storage and processing demands on the system.

4.8 Limitations of the Study

There are several limitations in the present study. This study included the use of a new visuo-spatial task. The task itself is somewhat unnatural since in order to match the sequence of demands in complex subject-object sentences, it was necessary to impose sequential processing onto visuo-spatial processing. This lack of ecological validity has two major implications. First, the novelty of the task meant that participants had to learn how to perform it; the "rules" of the task had to be explained, and in many cases, participants required extra coaching through the practice trials in order to achieve the criterion level to proceed to the test trials. The fact that the task required considerable practice before it could be adequately performed makes it different from on-line sentence processing, which requires no controlled processing. Furthermore, the visuo-spatial task is unlike sentence processing in that it does not serve to construct meaning; the visuo-spatial task was contrived in that there was no intrinsic goal to processing, whereas sentence processing is motivated by and leads to comprehension of a message. In a pilot of the present study, a second visuo-spatial task was created in which participants did construct meaning through performing the task. Participants viewed portions of pictures (e.g., an elephant, a helicopter) presented incrementally. The
picture fragments were not always presented right-side up, and so participants needed to mentally rotate the segments in order to determine whether the pieces created a cohesive picture. Unfortunately, no differences were found between SO and OS versions of the task, and so it was abandoned in favour of the current visuo-spatial task. Variations on the visuo-spatial operations task should be explored in future research because of its greater ecological validity. Second, because only one measure of visuo-spatial processing was used in the study, it is unknown how this visuo-spatial task correlates with established standardized tests of visuo-spatial processing (such as those used in Shah and Miyake, 1996). Thus, the extent to which this task is a "pure" measure of visuo-spatial processing is not established in this study. Indeed, many participants reported (or were observed) labelling the shapes on the frames as a verbal strategy for carrying out the visuo-spatial task. On the other hand, some participants reported using spatial strategies for responding to the comprehension question following each sentence (e.g., by visualizing where the noun phrase being questioned had appeared on the computer monitor). Given the interaction between verbal and visuo-spatial domains, it is unlikely that visuo-spatial tasks could entirely eliminate verbal processing.

Another limitation in the present study is the set of stimuli used for the syntactically ambiguous sentences. These stimuli were chosen in order to replicate previous studies which have employed the same stimuli (MacDonald et al., 1992; Kemtes & Kemper, 1997). However, more recent research has emphasized the importance of factors which guide the process of resolution of syntactic ambiguity, including the frequency with which a verb appears as past tense versus a past participle, and the "thematic fit" between noun-verb combinations (MacDonald et al., 1994; Pearlmutter & MacDonald, 1995; Trueswell & Kim, 1998; McRae, Spivey-Knowlton & Tanenhaus, 1998). The sentences used in this study did
not control for these factors, and so it is unknown what effects they may have had on performance. A small number of sentence sets (those using an inanimate noun as the first noun phrase) biased readers towards a reduced relative interpretation over a main verb interpretation. For example,

1. A yellow frisbee dropped from the roof onto the narrow driveway.
2. A yellow frisbee dropped from the roof landed in the ditch.

In this example, it seems much more likely that someone would drop the frisbee, rather than it dropping on its own (although the latter interpretation is possible). To my knowledge, no replication of previous studies of syntactical ambiguity resolution have employed stimuli sets which have controlled for these possible influencing factors.

4.9 Future Directions

The present study stimulates further inquiry into domains of memory and their relationship to sentence processing. The results of this study can be extended by using other sentence types (beyond subject-object and object-subject) and creating new "working memory" measures matched to the demands in those sentences. It may be fruitful to further investigate the possibility of creating a visuo-spatial processing task through which meaning is constructed and which is more natural than the task used in this study. It would also be interesting to include other more widely-used tests of visuo-spatial processing (see Shah & Miyake, 1996) to establish the validity of new visuo-spatial measures. An additional test of the robustness of the current results would be to replicate the present study using a different measure of sentence processing. It has been suggested that auditory processing of sentences is more automatic than reading sentences in a self-paced moving window presentation (Shapiro, personal communication, October, 2000). Therefore, methods such as cross-modal
interference (Shapiro, Swinney & Borsky, 1998) or auditory moving window (Ferreira &
Anes, 1994), in which participants listen to sentences, would extend the findings from this
study.

4.10 Conclusion

The results of this behavioural study provide support for the view of "working memory"
advanced by connectionist researchers. This study provides an example of how when
sequences of processing demands are matched across tasks, even tasks in different cognitive
domains can be good predictors of performance on sentence processing tasks. Thus, the
divisions between domains proposed by other models of working memory (Just & Carpenter,
1992; Caplan & Waters, 1999) may not be as well-founded as was previously thought. More
importantly, the entire concept of "resource pools" is drawn into question; a connectionist
account in which individual differences do not relate to the size of some undefined resource
pools, but rather are based in the architecture of the processor and experience, is supported
by these results. While there are many questions which remain to be answered, this study
provides a solid first step towards a new way of thinking about the nature of "resources"
underlying sentence processing ability.
REFERENCES


APPENDIX A

Syntactic Complexity Stimuli

Subject-Object Sentences

1. The teacher who the student admired won the award.
2. The pianist who the conductor praised attended the rehearsal.
3. The detective who the woman believed reported the evidence.
4. The senator who the journalist attacked admitted the error.
5. The artist who the curator slapped initiated the incident.
6. The florist who the employee helped addressed the envelopes.
7. The criminal who the lawyer doubted explained the situation.
8. The farmers who the politicians threatened disregarded their concerns.
9. The man who the performers impressed accepted the contract.
10. The musician who the surgeon kissed bought the flowers.
11. The carpenter who the plumber called provided excellent references.
12. The clerk who the officer thanked opened the door.
13. The girl who the boy disliked played on the swings.
14. The reporter who the anchorman observed commented on the situation.
15. The salesperson who the bookseller phoned wondered about the deal.
16. The bride who the organist surprised expected more traditional music.
17. The activists who the loggers feared promised to cause trouble.
18. The magician who the clown complimented stood on the stage.
19. The motorist who the cyclist cursed crossed the intersection illegally.
20. The researcher who the professor contacted suggested amendments to the paper.
21. The chief who the councilmen evaded demanded changes in the schedule.
22. The contractor who the homeowner insulted wanted to revise the plans.
23. The customer who the agent irritated waited for an answer impatiently.
24. The gardener who the runner greeted decided to take a break.
25. The technician who the operator reassured required the system upgrade immediately.
26. The actor who the director despised delayed the rehearsal.
27. The attorney who the witness trusted understood the problem.
28. The chef who the server ignored dropped the plate.
29. The illustrator who the author criticized abandoned the work.
30. The photographer who the model assisted needed a portfolio.
31. The secretary who the executive avoided requested the documents.
32. The scientist who the programmer commended detected the error.
33. The architect who the company challenged rejected the plans.
34. The foreigner who the shopkeeper confused asked the question.
35. The caterer who the waiter paid brought the food.
36. The assistant who the librarian annoyed worked long hours.
37. The administrator who the bureaucrat reprimanded wrote the report.
38. The player who the coach liked prepared for the game.
39. The client who the accountant admonished forgot about the claim.
40. The child who the sitter watched looked in the refrigerator.
41. The parents who the principal warned resisted the new idea.
42. The minister who the ambassador welcomed arrived despite many delays.
43. The nurse who the therapist congratulated was offered the promotion.
44. The hikers who the group paid maintained the forest trails.
45. The dancer who the instructor adored planned to choreograph the piece.
46. The debater who the opponent frightened neglected to prepare his notes.
47. The dentist who the repairman found advertised in the Yellow Pages.
48. The singer who the manager deceived dreamed about making a fortune.
49. The tourist who the bartender befriended hoped to find an apartment.
50. The doctor who the patient questioned wished he were somewhere else.

Object-Subject Sentences
1. The teacher admired the student who won the award.
2. The pianist praised the conductor who attended the rehearsal.
3. The detective believed the woman who reported the evidence.
4. The senator attacked the journalist who admitted the error.
5. The artist slapped the curator who initiated the incident.
6. The florist helped the employee who addressed the envelopes.
7. The criminal doubted the lawyer who explained the situation.
8. The farmers threatened the politicians who disregarded their concerns.
9. The man impressed the performers who accepted the contract.
10. The musician kissed the surgeon who bought the flowers.
11. The carpenter called the plumber who provided excellent references.
12. The clerk thanked the officer who opened the door.
13. The girl disliked the boy who played on the swings.
14. The reporter observed the anchorman who commented on the situation.
15. The salesperson phoned the bookseller who wondered about the deal.
16. The bride surprised the organist who expected more traditional music.
17. The activists feared the loggers who promised to cause trouble.
18. The magician complimented the clown who stood on stage.
19. The motorist cursed the cyclist who crossed the intersection illegally.
20. The researcher contacted the professor who suggested amendments to the paper.
21. The chief evaded the councilmen who demanded changes in the schedule.
22. The contractor insulted the homeowner who wanted to revise the plans.
23. The customer irritated the agent who waited for an answer impatiently.
24. The gardener greeted the runner who decided to take a break.
25. The technician reassured the operator who required the system upgrade immediately.
26. The actor despised the director who delayed the performance.
27. The attorney trusted the witness who understood the problem.
28. The chef ignored the server who dropped the plate.
29. The illustrator criticized the author who abandoned the work.
30. The photographer assisted the model who needed a portfolio.
31. The secretary avoided the executive who requested the documents.
32. The scientist commended the programmer who detected the error.
33. The architect challenged the company who rejected the plans.
34. The foreigner confused the shopkeeper who asked the question.
35. The caterer paid the waiter who brought the food.
36. The assistant annoyed the librarian who worked long hours.
37. The administrator reprimanded the bureaucrat who wrote the report.
38. The player liked the coach who prepared for the game.
39. The client admonished the accountant who forgot about the claim.
40. The child watched the sitter who looked in the refrigerator.
41. The parents warned the principal who resisted the new idea.
42. The minister welcomed the ambassador who arrived despite many delays.
43. The nurse congratulated the therapist who was offered the promotion.
44. The hikers paid the group who maintained the forest trails.
45. The dancer adored the instructor who planned to choreograph the piece.
46. The debater frightened the opponent who neglected to prepare his notes.
47. The dentist found the repairman who advertised in the Yellow Pages.
48. The singer deceived the manager who dreamed about making a fortune.
49. The tourist befriended the bartender who hoped to find an apartment.
50. The doctor questioned the patient who wished he were somewhere else.
APPENDIX B

Syntactic Ambiguity Stimuli

MacDonald et al. (1992) and Kemtes & Kemper (1997)

1. The experienced soldiers warned about the dangers before the midnight raid.
2. The experienced soldiers warned about the dangers conducted the midnight raid.
3. The experienced soldiers spoke about the dangers before the midnight raid.
4. The experienced soldiers who were told about the dangers conducted the midnight raid.

5. Several angry workers warned about low wages during the holiday season.
6. Several angry workers warned about low wages decided to file complaints.
7. Several angry workers spoke about low wages during the holiday season.
8. Several angry workers who were told about low wages decided to file complaints.

9. The soybean farmers warned about bad floods just before harvest time.
10. The soybean farmers warned about bad floods had no other crops.
11. The soybean farmers spoke about bad floods just before harvest time.
12. The soybean farmers who were told about bad floods had no other crops.

13. The ballet dancer asked about the routine before the rehearsal began.
14. The ballet dancer asked about the routine directed the other dancers.
15. The ballet dancer told about the routine before the rehearsal began.
16. The ballet dancer who had known about the routine directed the other dancers.

17. The young pupil asked about the questions before the exam began.
18. The young pupil asked about the questions failed to answer correctly.
19. The young pupil told about the questions before the exam began.
20. The young pupil who had known about the questions failed to answer correctly.

21. The soccer coach asked about the new player before the practice began.
22. The soccer coach asked about the new player said he missed practice.
23. The soccer coach told about the new player before the practice began.
24. The soccer coach who knew about the new player said he missed practice.

25. The efficient assistant phoned about the documents before the office opened.
26. The efficient assistant phoned about the documents prepared for the meeting.
27. The efficient assistant talked about the documents before the office opened.
28. The efficient assistant who was concerned about the documents prepared for the meeting.
29. The car salesperson phoned about the car before he decided to sell.
30. The car salesperson phoned about the car decided not to sell.
31. The car salesperson talked about the car before he decided to sell.
32. The car salesperson who was concerned about the car decided not to sell.
33. The forgetful child phoned about the meeting before school was over.
34. The forgetful child phoned about the meeting decided not to attend.
35. The forgetful child talked about the meeting before school was over.
36. The forgetful child who was concerned about the meeting decided not to attend.
37. The kitchen staff served in the cafeteria after the executives finished.
38. The kitchen staff served in the cafeteria soon got very sleepy.
39. The kitchen staff ate in the cafeteria after the executives finished.
40. The kitchen staff who were fed in the cafeteria soon got very sleepy.
41. The sunburned boys served the hot dogs at the football stadium.
42. The sunburned boys served the hotdogs got a stomach ache.
43. The sunburned boys ate the hot dogs at the football stadium.
44. The sunburned boys who were fed the hotdogs got a stomach ache.
45. The evil genie served the golden figs in the ancient temple.
46. The evil genie served the golden figs went into a trance.
47. The evil genie ate the golden figs in the ancient temple.
48. The evil genie who was fed the golden figs went into a trance.
49. The executive manager offered the job before he called his wife.
50. The executive manager offered the job despised the company's president.
51. The executive manager gave the job to the woman before he called his wife.
52. The executive manager who had found the job despised the company's president.
53. The skilled chef offered cooking tips before leaving the restaurant.
54. The skilled chef offered cooking tips decided to bake the dessert.
55. The skilled chef gave cooking tips before leaving the restaurant.
56. The skilled chef who had found cooking tips decided to bake the dessert.
57. The observant witness offered the false testimony to the examining prosecutor.
58. The observant witness offered the false testimony decided to tell the truth.
59. The observant witness gave the false testimony to the examining prosecutor.
60. The observant witness who had found the false testimony decided to tell the truth.
61. The calico cat washed in the alley after drinking the milk.
62. The calico cat washed in the alley ran into the street.
63. The calico cat cried in the alley after drinking the milk.
64. The calico cat who was bathed in the alley ran into the street.
65. The young children washed in the stream after the mothers left.
66. The young children washed in the stream splashed and shouted loudly.
67. The young children cried in the stream after the mothers left.
68. The young children who were bathed in the stream splashed and shouted loudly.

69. The sick child washed early every morning in the hospital room.
70. The sick child washed early every morning wanted her rubber duck.
71. The sick child cried early every morning in the hospital room.
72. The sick child who was bathed early every morning wanted her rubber duck.

73. The older kids taught all the routines for the spring recital.
74. The older kids taught all the routines danced in the recital.
75. The older kids learned all the routines for the spring recital.
76. The older kids who were shown all the routines danced in the recital.

77. The young technician taught the program from the thick manual.
78. The young technician taught the program caught on right away.
79. The young technician learned the program from the thick manual.
80. The young technician who was shown the program caught on right away.

81. The six volunteers taught the complicated procedure without very much trouble.
82. The six volunteers taught the complicated procedure became very good students.
83. The six volunteers learned the complicated procedure without very much trouble.
84. The six volunteers who were shown the complicated procedure became very good students.

85. A small dog pushed through the fence into the chicken coop.
86. A small dog pushed through the fence hurt his hind leg.
87. A small dog went through the fence into the chicken coop.
88. A small dog who was shoved through the fence hurt his hind leg.

89. The frightened kid pushed through the crowd to the front row.
90. The frightened kid pushed through the crowd looked for the nanny.
91. The frightened kid went through the crowd to the front row.
92. The frightened kid who was shoved through the crowd looked for the nanny.

93. An impatient shopper pushed through the doors to the sales racks.
94. An impatient shopper pushed through the doors complained to the manager.
95. An impatient shopper went through the doors to the sales racks.
96. An impatient shopper who was shoved through the doors complained to the manager.
97. The observant officers questioned about the homicide during the lengthy investigation.
98. The observant officers questioned about the homicide tolerated the lengthy investigation.
99. The observant officers inquired about the homicide during the lengthy investigation.
100. The observant officers who were interrogated about the homicide tolerated the lengthy investigation.
101. The annoyed teacher questioned about the misconduct after the students lied.
102. The annoyed teacher questioned about the misconduct knew the students lied.
103. The annoyed teacher inquired about the misconduct after the students lied.
104. The annoyed teacher who was interrogated about the misconduct knew the students lied.
105. The guilty banker questioned about the embezzlement during the police arrest.
106. The guilty banker questioned about the embezzlement observed the police arrest.
107. The guilty banker inquired about the embezzlement during the police arrest.
108. The guilty banker who was interrogated about the embezzlement observed the police arrest.
109. The eager children begged for the toys during the moving sale.
110. The eager children begged for the toys tolerated the moving sale.
111. The eager children pleaded for the toys during the moving sale.
112. The eager children who were asked for the toys tolerated the moving sale.
113. The graphic artist begged for an extension before the client meeting.
114. The graphic artist begged for an extension observed the client meeting.
115. The graphic artist pleaded for an extension before the client meeting.
116. The graphic artist who was asked for the extension observed the client meeting.
117. The irritated teachers begged about the students during the parent meetings.
118. The irritated teachers begged about the students avoided the parent meetings.
119. The irritated teachers pleaded about the students during the parent meetings.
120. The irritated teachers who were asked about the students avoided the parent meetings.
121. The brown sparrow watched on a branch pecked at an insect.
122. The brown sparrow watched on a branch high above the cat.
123. The brown sparrow stared on a branch high above the cat.
124. The brown sparrow who was seen on a branch pecked at an insect.
125. The young children watched in the hallway were following adults.
126. The young children watched in the hallway tolerated the adults argued.
127. The young children stared in the hallway while the adults argued.
128. The young children who were seen in the hallway were following adults.
129. The convicted criminal watched in the cell during the parole hearing.
130. The convicted criminal watched in the cell plotted a daring escape.
131. The convicted criminal stared in the cell during the parole hearing.
132. The convicted criminal who was seen in the cell plotted a daring escape.

133. A yellow frisbee dropped from the roof onto the narrow driveway.
134. A yellow frisbee dropped from the roof landed in the ditch.
135. A yellow frisbee fell from the roof onto the narrow driveway.
136. A yellow frisbee that was thrown from the roof landed in the ditch.

137. The large package dropped from the plane into the dark jungle.
138. The large package dropped from the plane hit several tall trees.
139. The large package fell from the plane into the dark jungle.
140. The large package that was thrown from the plane hit several tall trees.

141. Many small stones dropped from the cliff during the fierce storm.
142. Many small stones dropped from the cliff damaged passing cars below.
143. Many small stones fell from the cliff during the fierce storm.
144. Many small stones that were thrown from the cliff damaged passing cars below.

145. The silly boys entertained before the play until the teacher arrived.
146. The silly boys entertained before the play quickly left the auditorium.
147. The silly boys performed before the play until the teacher arrived.
148. The silly boys who were amused before the play quickly left the auditorium.

149. The teenage girls entertained in the hallway while the principle watched.
150. The teenage girls entertained in the hallway noticed the principle watched.
151. The teenage girls performed in the hallway while the principle watched.
152. The teenage girls who were amused in the hallway noticed the principle watched.

153. The thoughtless secretaries entertained on the balcony while the parade passed.
154. The thoughtless secretaries entertained on the balcony returned to their desks.
155. The thoughtless secretaries performed on the balcony while the parade passed.
156. The thoughtless secretaries who were amused on the balcony returned to their desks.

157. The engaging speaker asked about virtuous living during the community meeting.
158. The engaging speaker asked about virtuous living disliked the community meeting.
159. The engaging speaker inquired about virtuous living during the community meeting.
160. The engaging speaker who was informed about virtuous living disliked the community meeting.
161. The conscientious governor asked about the policy before the press conference.
162. The conscientious governor asked about the policy controlled the press conference.
163. The conscientious governor inquired about the policy before the press conference.
164. The conscientious governor who was informed about the policy controlled the press conference.
165. The skilled comedian asked about the audience during the comedy routine.
166. The skilled comedian asked about the audience performed the comedy routine.
167. The skilled comedian inquired about the audience during the comedy routine.
168. The skilled comedian who was informed about the audience performed the comedy routine.
169. The responsible homeowner observed for the newspaper during the chaotic riots.
170. The responsible homeowner observed for the newspaper relived the chaotic riots.
171. The responsible homeowner reported for the newspaper during the chaotic riots.
172. The responsible homeowner who was interviewed for the newspaper relived the chaotic riots.
173. The realty inspector observed for the worried family before the closing date.
174. The realty inspector observed for the worried family missed the closing date.
175. The realty inspector reported for the worried family before the closing date.
176. The realty inspector who was interviewed for the worried family missed the closing date.
177. The licensed doctor observed for the law team before the malpractice trial.
178. The licensed doctor observed for the law team jeopardized the malpractice trial.
179. The licensed doctor reported for the law team before the malpractice trial.
180. The licensed doctor who was interviewed for the law team jeopardized the malpractice trial.
181. The district judge queried about the perjured witness after the acquittal verdict.
182. The district judge queried about the perjured witness overturned the acquittal verdict.
183. The district judge inquired about the perjured witness after the acquittal verdict.
184. The district judge who was asked about the perjured witness overturned the acquittal verdict.
185. The research assistant queried about the new project before developing further plans.
186. The research assistant queried about the new project enjoyed developing further plans.
187. The research assistant inquired about the new project before developing further plans.
188. The research assistant who was asked about the new project enjoyed developing further plans.
189. The tennis coach queried about the champion players during the world finals.
190. The tennis coach queried about the champion players watched the world finals.
191. The tennis coach inquired about the champion players during the world finals.
192. The tennis coach who was asked about the champion players watched the world finals.
APPENDIX B (continued)

Comprehension Questions

MacDonald et al. (1992) and Kemtes & Kemper (1997)

1. Did the soldiers tell someone about the dangers? Y
2. Did the soldiers tell someone about the dangers? N
3. Did the soldiers tell someone about the dangers? Y
4. Did the soldiers tell someone about the dangers? N
5. Did the workers tell someone about the wages? Y
6. Did the workers tell someone about the wages? N
7. Did the workers tell someone about the wages? Y
8. Did the workers tell someone about the wages? N
9. Did the farmers tell someone about the floods? Y
10. Did the farmers tell someone about the floods? N
11. Did the farmers tell someone about the floods? Y
12. Did the farmers tell someone about the floods? N
13. Did the dancer ask someone about the routine? Y
14. Did the dancer ask someone about the routine? N
15. Did the dancer tell someone about the routine? Y
16. Did the dancer ask someone about the routine? N
17. Did the pupil ask someone about the questions? Y
18. Did the pupil ask someone about the questions? N
19. Did the pupil tell someone about the questions? Y
20. Did the pupil tell someone about the questions? N
21. Did the coach ask someone about the player? Y
22. Did the coach ask someone about the player? N
23. Did the coach tell someone about the player? Y
24. Did the coach ask someone about the player? N
25. Did the assistant call someone about the documents? Y
26. Did the assistant call someone about the documents? N
27. Did the assistant tell someone about the documents? Y
28. Was the assistant confident about the documents? N
29. Did the salesperson call someone about the car? Y
30. Did the salesperson call someone about the car? N
31. Did the salesperson tell someone about the car? Y
32. Was the salesperson confident about the car? N
33. Did the child call someone about the meeting? Y
34. Did the child call someone about the meeting? N
35. Did the child tell someone about the meeting? Y
36. Was the child confident about the meeting? N
37. Did the staff serve someone in the cafeteria? Y
38. Did the staff serve someone in the cafeteria? N
39. Did the staff eat in the cafeteria? Y
40. Did the staff serve someone in the cafeteria? N
41. Did the boys serve hotdogs? Y
42. Did the boys serve hotdogs? N
43. Did the boys eat hotdogs? Y
44. Did the boys serve hotdogs? N
45. Did the genie serve figs? Y
46. Did the genie serve figs? N
47. Did the genie eat figs? Y
48. Did the genie serve figs? N
49. Did the executive offer the job? Y
50. Did the executive offer the job? N
51. Did the executive offer the job? Y
52. Did the executive offer the job? N
53. Did the chef offer cooking tips? Y
54. Did the chef offer cooking tips? N
55. Did the chef offer cooking tips? Y
56. Did the chef offer cooking tips? N
57. Did the witness offer false testimony? Y
58. Did the witness offer false testimony? N
59. Did the witness offer false testimony? Y
60. Did the witness offer false testimony? N
61. Did the cat wash itself in the alley? Y
62. Did the cat wash itself in the alley? N
63. Did the cat cry in the alley? Y
64. Did the cat wash itself in the alley? N
65. Did the children wash themselves in the stream? Y
66. Did the children wash themselves in the stream? N
67. Did the children cry in the stream? Y
68. Did the children wash themselves in the stream? N
69. Did the child wash herself early every morning? Y
70. Did the child wash herself early every morning? N
71. Did the child wash cry early every morning? Y
72. Did the child wash herself early every morning? N
73. Did the kids teach the routines? Y
74. Did the kids teach the routines? N
75. Did the kids teach the routines? Y
76. Did the kids teach the routines? N
77. Did the technician teach the program? Y
78. Did the technician teach the program? N
79. Did the technician learn the program? Y
80. Did the technician teach the program? N
81. Did the volunteers teach the procedure? Y
82. Did the volunteers teach the procedure? N
83. Did the volunteers learn the procedure? Y
84. Did the volunteers teach the procedure? N
85. Did someone push the dog through the fence? N
86. Did someone push the dog through the fence? Y
87. Did someone push the dog through the fence? N
88. Did someone push the dog through the fence? Y
89. Did someone push the kid through the crowd? N
90. Did someone push the kid through the crowd? Y
91. Did someone push the kid through the crowd? N
92. Did someone push the kid through the crowd? Y
93. Did the shopper push through the doors? Y
94. Did the shopper push through the doors? N
95. Did the shopper push through the doors? Y
96. Did the shopper push through the doors? N
97. Did the officers ask someone about the homicide? Y
98. Did the officers ask someone about the homicide? N
99. Did the officers ask someone about the homicide? Y
100. Did the officers ask someone about the homicide? N
101. Did the teacher question someone about the misconduct? Y
102. Did the teacher question someone about the misconduct? N
103. Did the teacher question someone about the misconduct? Y
104. Did the teacher question someone about the misconduct? N
105. Did the banker ask someone about the arrest? Y
106. Did the banker ask someone about the arrest? N
107. Did the banker ask someone about the arrest? Y
108. Did the banker ask someone about the arrest? N
109. Did the children beg someone for the toys? Y
110. Did the children beg someone for the toys? N
111. Did the children beg someone for the toys? Y
112. Did the children beg someone for the toys? N
113. Did the artist beg someone for an extension? Y
114. Did the artist beg someone for an extension? N
115. Did the artist beg someone for an extension? Y
116. Did the artist ask someone for an extension? N
117. Did the teachers beg someone about the students? Y
118. Did the teachers beg someone about the students? N
119. Did the teachers beg someone about the students? Y
120. Did the teachers ask someone about the students? N
121. Did the sparrow watch something from a branch? Y
122. Did the sparrow watch something from a branch? N
123. Did the sparrow watch something from a branch? Y
124. Did the sparrow watch something from a branch? N
125. Did the children watch someone in the hallway? N
126. Did the children watch someone in the hallway? Y
127. Did the children watch someone in the hallway? Y
128. Did the children watch someone in the hallway? N
129. Did the criminal watch someone from the cell? Y
130. Did the criminal watch someone from the cell? N
131. Did the criminal watch someone from the cell? Y
132. Did the criminal watch someone from the cell? N
133. Did the frisbee fall from the roof? Y
134. Did the frisbee fall from the roof? N
135. Did the frisbee fall from the roof? Y
136. Did the frisbee fall from the roof? N
137. Did the package fall from the plane? Y
138. Did the package fall from the plane? N
139. Did the package fall from the plane? Y
140. Did the package fall from the plane? N
141. Did the stones fall from the cliff? Y
142. Did the stones fall from the cliff? N
143. Did the stones fall from the cliff? Y
144. Did the stones fall from the cliff? N
145. Did the boys entertain before the play? Y
146. Did the boys entertain before the play? N
147. Did the boys entertain before the play? Y
148. Did the boys entertain before the play? N
149. Did the girls entertain in the hallway? Y
150. Did the girls entertain in the hallway? N
151. Did the girls entertain in the hallway? Y
152. Did the girls entertain in the hallway? N
153. Did the secretaries entertain someone on the balcony? Y
154. Did the secretaries entertain someone on the balcony? N
155. Did the secretaries entertain someone on the balcony? Y
156. Did the secretaries entertain someone on the balcony? N
157. Did the speaker ask about virtuous living? Y
158. Did the speaker ask about virtuous living? N
159. Did the speaker tell someone about virtuous living? Y
160. Was the speaker unaware of virtuous living? N
161. Did the governor ask about the policy? Y
162. Did the governor ask about the policy? N
163. Did the governor tell someone about the policy? Y
164. Was the governor unaware of the policy? N
165. Did the comedian ask about the audience? Y
166. Did the comedian ask about the audience? N
167. Did the comedian tell someone about the audience? Y
168. Was the comedian unaware of the audience? N
169. Did the homeowner report for the newspaper? Y
170. Did the homeowner report for the newspaper? N
171. Did the homeowner report for the newspaper? Y
172. Did the homeowner write a report for the newspaper? N
173. Did the inspector report for the family? Y
174. Did the inspector report for the family? N
175. Did the inspector report for the family? Y
176. Did the inspector write a report for the family? N
177. Did the doctor report for the law team? Y
178. Did the doctor report for the law team? N
179. Did the doctor report for the law team? Y
180. Did the doctor write a report for the law team? N
181. Did the judge ask about the witness? Y
182. Did the judge ask about the witness? N
183. Did the judge ask about the witness? Y
184. Did the judge ask about the witness? N
185. Did the assistant ask about the project? Y
186. Did the assistant ask about the project? N
187. Did the assistant ask about the project? Y
188. Did the assistant ask about the project? N
189. Did the coach ask about the players? Y
190. Did the coach ask about the players? N
191. Did the coach ask about the players? Y
192. Did the coach ask about the players? N
1  Did someone tell the soldiers about the dangers? N
2  Did someone tell the soldiers about the dangers? Y
3  Did someone tell the soldiers about the dangers? N
4  Did someone tell the soldiers about the dangers? Y
5  Did someone tell the workers about the wages? N
6  Did someone tell the workers about the wages? Y
7  Did someone tell the workers about the wages? N
8  Did someone tell the workers about the wages? Y
9  Did someone tell the farmers about the floods? N
10 Did someone tell the farmers about the floods? Y
11 Did someone tell the farmers about the floods? N
12 Did someone tell the farmers about the floods? Y
13 Did someone ask the dancer about the routine? N
14 Did someone ask the dancer about the routine? Y
15 Did someone tell the dancer about the routine? N
16 Did someone tell the dancer about the routine? Y
17 Did someone ask the pupil about the questions? N
18 Did someone ask the pupil about the questions? Y
19 Did someone tell the pupil about the questions? N
20 Did someone tell the pupil about the questions? Y
21 Did someone ask the coach about the player? N
22 Did someone ask the coach about the player? Y
23 Did someone tell the coach about the player? N
24 Did someone tell the coach about the player? Y
25 Did someone call the assistant about the documents? N
26 Did someone call the assistant about the documents? Y
27 Did someone tell the assistant about the documents? N
28 Was the assistant concerned about the documents? Y
29 Did someone call the salesperson about the car? N
30 Did someone call the salesperson about the car? Y
31 Did someone tell the salesperson about the car? N
32 Was the salesperson concerned about the car? Y
33. Did someone call the child about the meeting? N
34. Did someone call the child about the meeting? Y
35. Did someone tell the child about the meeting? N
36. Was the child concerned about the meeting? Y
37. Did someone serve the staff in the cafeteria? N
38. Did someone serve the staff in the cafeteria? Y
39. Did someone serve the staff in the cafeteria? N
40. Did someone serve the staff in the cafeteria? Y
41. Did someone serve hotdogs to the boys? N
42. Did someone serve hotdogs to the boys? Y
43. Did someone serve hotdogs to the boys? N
44. Did someone serve hotdogs to the boys? Y
45. Did someone serve the genie figs? N
46. Did someone serve the genie figs? Y
47. Did someone serve the genie figs? N
48. Did someone serve the genie figs? Y
49. Did someone offer the job to the executive? N
50. Did someone offer the job to the executive? Y
51. Did someone offer the job to the executive? N
52. Did someone offer the job to the executive? Y
53. Did someone offer the chef cooking tips? N
54. Did someone offer the chef cooking tips? Y
55. Did someone offer the chef cooking tips? N
56. Did someone offer the chef cooking tips? Y
57. Did someone offer the witness false testimony? N
58. Did someone offer the witness false testimony? Y
59. Did someone offer the witness false testimony? N
60. Did someone offer the witness false testimony? Y
61. Did someone wash the cat in the alley? N
62. Did someone wash the cat in the alley? Y
63. Did the cat sleep in the alley? N
64. Did someone wash the cat in the alley? Y
65. Did someone was the children in the stream? N
66. Did someone was the children in the stream? Y
67. Were the children happy? N
68. Did someone was the children in the stream? Y
69. Did someone wash the child early every morning? N
70. Did someone wash the child early every morning? Y
71. Was the child happy? N
72. Did someone wash the child early every morning? Y
73. Did someone teach the kids the routines? N
74. Did someone teach the kids the routines? Y
75. Did the kids learn the routines? N
76. Did someone teach the kids the routines? Y
77. Did someone teach the technician the program? N
78. Did someone teach the technician the program? Y
137

79. Did the technician learn the program? N
80. Did someone teach the technician the program? Y
81. Did someone teach the volunteers the procedure? N
82. Did someone teach the volunteers the procedure? Y
83. Did the volunteers learn the procedure? N
84. Did someone teach the volunteers the procedure? Y
85. Did the dog push through the fence? Y
86. Did the dog push through the fence? N
87. Did the dog push through the fence? Y
88. Did the dog push through the fence? N
89. Did the kid push through the crowd? Y
90. Did the kid push through the crowd? N
91. Did the kid push through the crowd? Y
92. Did the kid push through the crowd? N
93. Did someone push the shopper through the doors? N
94. Did someone push the shopper through the doors? Y
95. Did someone push the shopper through the doors? N
96. Did someone push the shopper through the doors? Y
97. Did someone ask the officers about the homicide? N
98. Did someone ask the officers about the homicide? Y
99. Did someone ask the officers about the homicide? N
100. Did someone ask the officers about the homicide? Y
101. Did someone question the teacher about the misconduct? N
102. Did someone question the teacher about the misconduct? Y
103. Did someone question the teacher about the misconduct? N
104. Did someone question the teacher about the misconduct? Y
105. Did someone ask the banker about the arrest? N
106. Did someone ask the banker about the arrest? Y
107. Did someone ask the banker about the arrest? N
108. Did someone ask the banker about the arrest? Y
109. Did someone beg the children for the toys? N
110. Did someone beg the children for the toys? Y
111. Did someone beg the children for the toys? N
112. Did someone ask the children for the toys? Y
113. Did someone beg the artist for an extension? N
114. Did someone beg the artist for an extension? Y
115. Did someone beg the artist for an extension? N
116. Did someone ask the artist for an extension? Y
117. Did someone beg the teachers about the students? N
118. Did someone beg the teachers about the students? Y
119. Did someone beg the teachers about the students? N
120. Did someone ask the teachers about the students? Y
121. Did something watch the sparrow on a branch? N
122. Did something watch the sparrow on a branch? Y
123. Did something watch the sparrow on a branch? N
124. Did something watch the sparrow on a branch? Y
125. Did someone watch the children in the hallway? Y
126. Did someone watch the children in the hallway? N
127. Did someone watch the children in the hallway? N
128. Did someone watch the children in the hallway? Y
129. Did someone watch the criminal in the cell? N
130. Did someone watch the criminal in the cell? Y
131. Did someone watch the criminal in the cell? N
132. Did someone watch the criminal in the cell? Y
133. Did someone throw the frisbee from the roof? N
134. Did someone throw the frisbee from the roof? Y
135. Did someone throw the frisbee from the roof? N
136. Did someone throw the frisbee from the roof? Y
137. Did someone throw the package from the plane? N
138. Did someone throw the package from the plane? Y
139. Did someone throw the package from the plane? N
140. Did someone throw the package from the plane? Y
141. Did someone throw the stones from the cliff? N
142. Did someone throw the stones from the cliff? Y
143. Did someone throw the stones from the cliff? N
144. Did someone throw the stones from the cliff? Y
145. Did someone entertain the boys before the play? N
146. Did someone entertain the boys before the play? Y
147. Did someone entertain the boys before the play? N
148. Did someone amuse the boys before the play? Y
149. Did someone entertain the girls in the hallway? N
150. Did someone entertain the girls in the hallway? Y
151. Did someone entertain the girls in the hallway? N
152. Did someone amuse the girls in the hallway? Y
153. Did someone entertain the secretaries on the balcony? N
154. Did someone entertain the secretaries on the balcony? Y
155. Did someone entertain the secretaries on the balcony? N
156. Did someone amuse the secretaries on the balcony? Y
157. Did someone ask the speaker about virtuous living? N
158. Did someone ask the speaker about virtuous living? Y
159. Did someone ask the speaker about virtuous living? N
160. Did the speaker find out about virtuous living? Y
161. Did someone ask the governor about the policy? N
162. Did someone ask the governor about the policy? Y
163. Did someone ask the governor about the policy? N
164. Did the governor find out about the policy? Y
165. Did someone ask the comedian about the audience? N
166. Did someone ask the comedian about the audience? Y
167. Did someone ask the comedian about the audience? N
168. Did the comedian find out about the audience? Y
169. Did the newspaper interview the homeowner? N
170. Did the newspaper interview the homeowner? Y
<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did the newspaper interview the homeowner?</td>
<td>N</td>
</tr>
<tr>
<td>Did the newspaper interview the homeowner?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone interview the inspector?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone interview the inspector?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone interview the inspector?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone interview the inspector?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone interview the doctor?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone interview the doctor?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone interview the doctor?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone interview the doctor?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone ask the judge about the witness?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone ask the judge about the witness?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone ask the judge about the witness?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone ask the judge about the witness?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone ask the assistant about the project?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone ask the assistant about the project?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone ask the assistant about the project?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone ask the assistant about the project?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone ask the coach about the players?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone ask the coach about the players?</td>
<td>Y</td>
</tr>
<tr>
<td>Did someone ask the coach about the players?</td>
<td>N</td>
</tr>
<tr>
<td>Did someone ask the coach about the players?</td>
<td>Y</td>
</tr>
</tbody>
</table>
Span 2

In a flash of fatigue and fantasy, he saw a fat Indian sitting beside a campfire. The lieutenant sat beside the man with the walkie-talkie and stared at the muddy ground.

I will not shock my readers with a description of the cool-blooded butchery that followed. The courses are designed as much for professional engineers as for amateur enthusiasts.

The taxi turned up Michigan Avenue, where they had a clear view of the lake. The words of human love have been used by the saints to describe their vision of God.

It was shortly after this that an unusual pressure of business called me into town. He pursued this theme, still pretending to seek for information to quiet his own doubts.

I was so surprised at this unaccountable apparition, that I was speechless for a while. When at last his eyes opened, there was no gleam of triumph, no shade of anger.

Span 3

Filled with these dreary forebodings, I fearfully opened the heavy wooden door. I'm not certain what went wrong but I think it was my cruel and bad temper. I imagine that you have a shrewd suspicion of the object of my early visit.

I turned my memories over at random like pictures in a photograph album. Sometimes I get so tired of trying to convince him that I love him and shall forever. The girl hesitated for a moment to taste the onions because her husband hated the smell.

It was your belief in the significance of my suffering that kept me going. When in trouble, children naturally hope for a miraculous intervention by a superhuman. With shocked amazement and appalled fascination Marion looked at the pictures.

There are days when the city where I live wakes in the morning with a strange look. We boys wanted to warn them, but we backed down when it came to the pinch. He stood there at the edge of the crowd while they were singing, and he looked bitter.

What would come after this day would be inconceivably different, would be real life. John became annoyed with Karen's bad habits of biting her hails and chewing gum. Due to his gross inadequacies, his position as director was terminated abruptly.
It is possible, of course, that life did not arise on the earth at all.
The poor lady was thoroughly persuaded that she was not long to survive this vision.
After all he had not gone far, and some of his walking had been circular.
The announcement of it would resound throughout the world, penetrate to the remotest land.

To do so in directions that are adaptive for mankind would be a realistic objective.
Slicing it out carefully with his knife, he folded it without creasing the face.
He laughed sarcastically and looked as if he could have poisoned me for my errors.
He tolerated another intrusion and thought himself a paragon of patience for doing so.

The reader may suppose that I had other motives, besides the desire to escape the law.
On the desk where she wrote her letters was a clutter of objects coated in dust.
He stuffed his denim jacket into his pants and fastened the stiff, new snaps securely.
He had an odd elongated skull which sat on his shoulders like a pear on a dish.

His imagination had so abstracted him that his name was called twice before he answered.
The basic characteristic of the heroes in the preceding stories is their sensitivity.
He listened carefully because he had the weird impression that he knew the voices.
He had patronized her when she was a schoolgirl and teased her when she was a student.

The rain and howling wind kept beating against the rattling window panes.
He covered his heart with both hands to keep anyone from hearing the noise it made.
The stories all deal with a middle-aged protagonist who attempts to withdraw from society.
Without tension there could be no balance either in nature or in mechanical design.

I wish there existed someone to whom I could say that I felt very sorry.
Here, as elsewhere, the empirical patterns are important and abundantly documented.
The intervals of silence grew progressively longer; the delays became very maddening.
Two or three substantial pieces of wood smoldered on the hearth, for the night was cold.
I imagined that he had been thinking things over while the secretary was with us.

There was still more than an hour before breakfast, and the house was silent and asleep.
He leant on the parapet of the bridge and the two policemen watched him from a distance.
These splendid melancholy eyes were turned upon me from the mirror with a haughty stare.
He sometimes considered suicide but the thought was too oppressive to remain in his mind.
And now that a man had died some unimaginably different state of affairs must come to be.
When I got to the big tobacco field I saw that it had not suffered much.
The products of digital electronics will play an important role in your future.
One problem with this explanation is that there appears to be no defense against cheating.
Sometimes the scapegoat is an outsider who has been taken into the community.
I should not be able to make any one understand how exciting it all was.

A small oil lamp burned on the floor and two men crouched against the wall, watching me.
The sound of an approaching train woke him, and he started to his feet.
The boisterous laughter of the children was disturbing to the aged in the building.
In comparison to his earlier works, the musician had developed a unique, enthralling style.
The entire construction crew decided to lengthen their work day in order to have lunch.

The smokers were asked to refrain from their habit until the end of the production.
All students that passed the test were exempt from any further seminars that semester.
Despite the unusually cold weather, the campers continued their canoe trip.
The young business executive was determined to develop his housing projects within the year.
In order to postpone the business trip, he cancelled his engagements for the week.

Span 6

The incorrigible child was punished brutally for his lack of respect for his elders.
The brilliant trial attorney dazzled the jury with his astute knowledge of the case.
I found the keynote speaker incredibly boring, inarticulate and not well read.
The devastating effects of the flood were not fully realized until months later.
In a moment of complete spontaneity, she developed a thesis for her paper.
At the conclusion of the musicians' performance, the enthusiastic crowd applauded.

Her mother nagged incessantly about her lack of concern for the welfare of the children.
Circumstantial evidence indicated that there was a conspiracy to eliminate him.
Without any hesitation, he plunged into the difficult mathematics assignment blindly.
To determine the effects of the medication, the doctor hospitalized his patient.
The lumbermen worked long hours in order to obtain the necessary amount of wood.
They attended the theater habitually except for circumstances beyond their control.

The old lady talked to her new neighbor on her weekly walk from church.
After passing all the exams, the class celebrated for an entire week without resting.
The entire town arrived to see the appearance of the controversial political candidate.
The weather was very unpredictable that summer so no one made plans too far in advance.
According to the results from the survey, Robert Redford is the most liked Hollywood star.
Jane's relatives had decided that her gentleman friend was not of high status.
We felt that this young troupe deserved an opportunity to perform before an American audience.
If you want to be understood quickly, do not waste time by using negatives.
Their aim is to tell educators and parents how to identify symptoms of drug abuse.
Then when his son suffered a final relapse, the man found himself unable to grieve.
At the age of eight she was admitted to the hospital, apparently unconscious.
A large number of well-educated Americans accept myths and misconceptions about welfare.

One thing I've noticed in my mail is how women denigrate the job of housewife.
A newborn baby does not have much in the way of social grace.
You know that we don't overnight become the powerful adult of our childhood dreams.
Caring for contact lenses is really no more taxing than good dental hygiene.
Her waters are polluted, her soil is becoming dust, some of her trees are poisoned.
It's no accident that in recent years I've changed my style of teaching radically.
APPENDIX D

Waters and Caplan (1996a, 1996b) Reading Span Materials

Span 2

It was the spectators that watched the play.
It was the child that bit into the fruit.

It was the drum that beat the child.
It was the disc jockey that broke the tape.

It was the baby that clenched the cup.
It was the pizza that bit into the boy.

It was the ice cream that requested the child.
It was the chairs that counted the bride.

It was the statue that defaced the kids.
It was the shipper that mailed the box.

Span 3

It was the puzzle that intrigued the child.
It was the woman that fascinated the art.
It was the song that played the tape.

It was the woman that inherited the house.
It was the bumpy road that slowed down the cars.
It was the ticket that scratched the man.

It was the pie that cut the cook.
It was the ocean that swallowed up the boat.
It was the man that appealed to the clothes.

It was the marbles that lost the child.
It was the woman that cherished the note.
It was the puzzle that solved the girl.

It was the woman that delighted the gift.
It was the cook that heated the pot.
It was the rope that tripped the horse.
Span 4

It was the woman that nibbled on the food.
It was the cookie that made the girl.
It was the thief that tempted the house.
It was the weapon that incriminated the man.

It was the candy that perked up the child.
It was the elephant that knocked over the gate.
It was the critic that disappointed the play.
It was the radio that annoyed the class.

It was the baby that suffocated the sheet.
It was the gloves that held up the kid.
It was the souvenir that impressed the man.
It was the fleeing man that hindered the tree.

It was the basement that flooded the rain.
It was the candles that started the fire.
It was the girl that nourished the milk.
It was the rats that finished off the traps.

It was the connoisseur that enjoyed the wine.
It was the martini that relaxed the man.
It was the hot chocolate that enjoyed the child.
It was the airplane that thrilled the boy.

Span 5

It was the dollhouse that amazed the child.
It was the bracelet that pleased the girl.
It was the bullet that finished off the horse.
It was the gangster that broke into the store.
It was the furniture that polished the wife.

It was the driver that frustrated the lane.
It was the swimmer that guided the lines.
It was the pedestrian that tripped over the branch.
It was the cup that glued the man.
It was the painting that inspired the pope.

It was the mess that bothered the maid.
It was the television that distracted the child.
It was the man that annoyed the light.
It was the tailor that displeased the cloth.
It was the mirror that excited the cat.

It was the maid that polished the sink.
It was the housewife that angered the cans.
It was the policeman that provoked the gun.
It was the room that cramped the guests.
It was the coffee that woke up the man.

It was the gangster that cut the knife.
It was the child that bewildered the game.
It was the people that excited the play.
It was the inspector that rejected the place.
It was the tenant that irritated the door.

Span 6

It was the director that began the play.
It was the air conditioner that installed the man.
It was the bride that terrified the dress.
It was the check that overlooked the bank.
It was the timer that alerted the cook.
It was the symphony that delighted the queen.

It was the newspapers that accumulated the maid.
It was the student that took the test.
It was the jogger that lost the hat.
It was the man that pleased the tie.
It was the burglar that tripped the hole.
It was the collector that pleased the chair.

It was the people that hurt the bomb.
It was the nightmare that frightened the child.
It was the vacation that enjoyed the king.
It was the vacationers that stranded the flood.
It was the housewife that dropped the key.
It was the assassin that shot the pope.

It was the throne that was on the lord.
It was the shopper that enticed the clothes.
It was their testimony that swayed the court.
It was the banker that issued the loan.
It was the thief that tempted the rings.
It was the fog that crashed the plane.
It was the official that seized the drugs.
It was the handkerchief that gripped the girl.
It was the woman that treasured the glass.
It was the chair that sat on the child.
It was the ship that transported the food.
It was the woman that chose the dress.
APPENDIX E

Mathematical Operations Stimuli

Subject-Object Constructions

1. \( 52 \div (1 + 3) + 2 = 15 \)
2. \( 50 \div (4 + 1) + 4 = 14 \)
3. \( 56 \div (2 + 2) - 2 = 12 \)
4. \( 48 \div (1 + 2) - 3 = 13 \)
5. \( 42 \div (1 + 2) - 1 = 13 \)
6. \( 60 \div (8 - 3) + 8 = 20 \)
7. \( 44 \div (10 - 6) + 7 = 18 \)
8. \( 45 \div (7 - 4) - 2 = 13 \)
9. \( 55 \div (10 - 5) - 1 = 10 \)
10. \( 48 \div (6 - 2) - 1 = 11 \)

11. \( 8 \times (1 + 1) + 4 = 20 \)
12. \( 4 \times (2 + 1) + 6 = 18 \)
13. \( 3 \times (2 + 4) + 1 = 19 \)
14. \( 5 \times (1 + 3) - 9 = 11 \)
15. \( 4 \times (3 + 1) - 5 = 11 \)
16. \( 3 \times (9 - 4) + 4 = 19 \)
17. \( 2 \times (10 - 3) + 3 = 17 \)
18. \( 7 \times (8 - 6) + 4 = 18 \)
19. \( 6 \times (5 - 2) - 6 = 12 \)
20. \( 5 \times (7 - 4) + 1 = 16 \)
Object-Subject Constructions

1. $50 \div 5 + 4 + 5 = 19$
2. $44 \div 4 + 3 + 6 = 20$
3. $60 \div 5 + 6 - 5 = 13$
4. $48 \div 4 + 7 - 7 = 12$
5. $55 \div 5 + 9 - 6 = 14$
6. $48 \div 3 - 3 + 6 = 19$
7. $40 \div 2 - 9 + 6 = 17$
8. $56 \div 4 - 1 - 1 = 12$
9. $45 \div 3 - 4 - 1 = 10$
10. $42 \div 3 - 2 - 2 = 10$

11. $4 \times 3 + 6 + 2 = 20$
12. $5 \times 2 + 7 + 1 = 18$
13. $2 \times 7 + 2 + 3 = 19$
14. $3 \times 6 + 2 - 8 = 12$
15. $2 \times 9 + 1 - 5 = 14$
16. $4 \times 5 - 8 + 3 = 15$
17. $8 \times 2 - 3 + 4 = 17$
18. $3 \times 5 - 1 + 2 = 16$
19. $5 \times 4 - 5 - 4 = 11$
20. $4 \times 4 - 2 - 1 = 13$
APPENDIX F

Examples of the visuo-spatial stimuli

Subject-Object Condition

Location

Orientation

Shading

Screen 1 | Screen 2 | Screen 3

The correct answers are: Location, 1; Orientation, 2; and Shading, 1.
Object-subject Condition

Location

Orientation

Shading

Screen 1  Screen 2  Screen 3

The correct answers are: Location, 1; Orientation, 2; Shading, 2.

Note. The size of the examples has been reduced to fit the page; the actual stimuli appeared on the screen in $4 \times 5$ inch frames.


APPENDIX G

Participants' Scores on Standard Working Memory Measures

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M = 3.1$</td>
<td>$M = 4.4$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD = 0.8$</td>
<td>$SD = 1.3$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4.0</td>
<td>4.0</td>
<td>.187</td>
</tr>
<tr>
<td>2</td>
<td>3.0</td>
<td>5.0</td>
<td>.502</td>
</tr>
<tr>
<td>3</td>
<td>3.0</td>
<td>3.0</td>
<td>-.933</td>
</tr>
<tr>
<td>4</td>
<td>3.0</td>
<td>4.5</td>
<td>.124</td>
</tr>
<tr>
<td>5</td>
<td>4.5</td>
<td>5.5</td>
<td>---</td>
</tr>
<tr>
<td>6</td>
<td>2.0</td>
<td>5.5</td>
<td>.855</td>
</tr>
<tr>
<td>7</td>
<td>3.0</td>
<td>3.5</td>
<td>.303</td>
</tr>
<tr>
<td>8</td>
<td>4.0</td>
<td>4.0</td>
<td>-.310</td>
</tr>
<tr>
<td>9</td>
<td>4.0</td>
<td>5.5</td>
<td>-.722</td>
</tr>
<tr>
<td>10</td>
<td>2.0</td>
<td>5.0</td>
<td>1.022</td>
</tr>
<tr>
<td>11</td>
<td>3.0</td>
<td>3.0</td>
<td>.346</td>
</tr>
<tr>
<td>12</td>
<td>3.0</td>
<td>1.5</td>
<td>-.052</td>
</tr>
<tr>
<td>13</td>
<td>2.5</td>
<td>4.5</td>
<td>.378</td>
</tr>
<tr>
<td>14</td>
<td>2.0</td>
<td>2.5</td>
<td>-1.474</td>
</tr>
<tr>
<td>15</td>
<td>3.0</td>
<td>5.0</td>
<td>-.187</td>
</tr>
<tr>
<td>16</td>
<td>3.0</td>
<td>5.5</td>
<td>.001</td>
</tr>
<tr>
<td>17</td>
<td>3.0</td>
<td>2.0</td>
<td>-.723</td>
</tr>
<tr>
<td>18</td>
<td>2.5</td>
<td>5.0</td>
<td>.284</td>
</tr>
<tr>
<td>19</td>
<td>3.5</td>
<td>6.0</td>
<td>.351</td>
</tr>
<tr>
<td>20</td>
<td>3.0</td>
<td>4.0</td>
<td>-.984</td>
</tr>
<tr>
<td>21</td>
<td>3.0</td>
<td>1.5</td>
<td>-.705</td>
</tr>
<tr>
<td>22</td>
<td>3.0</td>
<td>5.5</td>
<td>-.068</td>
</tr>
<tr>
<td>23</td>
<td>2.0</td>
<td>5.0</td>
<td>1.182</td>
</tr>
<tr>
<td>24</td>
<td>3.0</td>
<td>5.5</td>
<td>.318</td>
</tr>
<tr>
<td>25</td>
<td>3.0</td>
<td>5.0</td>
<td>.145</td>
</tr>
<tr>
<td>26</td>
<td>5.5</td>
<td>5.0</td>
<td>-.071</td>
</tr>
<tr>
<td>27</td>
<td>2.5</td>
<td>4.0</td>
<td>.086</td>
</tr>
<tr>
<td>28</td>
<td>3.5</td>
<td>5.5</td>
<td>.210</td>
</tr>
<tr>
<td>29</td>
<td>2.5</td>
<td>5.0</td>
<td>.011</td>
</tr>
<tr>
<td>30</td>
<td>3.0</td>
<td>4.5</td>
<td>-.075</td>
</tr>
</tbody>
</table>

Note. Reaction time data on the Waters and Caplan reading span task were not collected for participant number five due to a technical error. A composite score could therefore not be calculated for this participant.
APPENDIX H

Participants’ Mean Scores on Mathematical and Visuo-spatial Operations Working Memory Tasks (in msec)

<table>
<thead>
<tr>
<th>Subject no.</th>
<th>Critical regions</th>
<th>Aggregate scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Math</td>
<td>Visuo-spatial</td>
</tr>
<tr>
<td></td>
<td>$M = 2011$</td>
<td>$M = 3362$</td>
</tr>
<tr>
<td></td>
<td>$SD = 572$</td>
<td>$SD = 1174$</td>
</tr>
<tr>
<td>1</td>
<td>1739</td>
<td>2217</td>
</tr>
<tr>
<td>2</td>
<td>3426</td>
<td>3818</td>
</tr>
<tr>
<td>3</td>
<td>2735</td>
<td>2930</td>
</tr>
<tr>
<td>4</td>
<td>1933</td>
<td>2461</td>
</tr>
<tr>
<td>5</td>
<td>1594</td>
<td>2534</td>
</tr>
<tr>
<td>6</td>
<td>1600</td>
<td>2363</td>
</tr>
<tr>
<td>7</td>
<td>1569</td>
<td>2254</td>
</tr>
<tr>
<td>8</td>
<td>1880</td>
<td>2396</td>
</tr>
<tr>
<td>9</td>
<td>1688</td>
<td>1622</td>
</tr>
<tr>
<td>10</td>
<td>2527</td>
<td>4886</td>
</tr>
<tr>
<td>11</td>
<td>2756</td>
<td>5593</td>
</tr>
<tr>
<td>12</td>
<td>2749</td>
<td>3921</td>
</tr>
<tr>
<td>13</td>
<td>2312</td>
<td>3317</td>
</tr>
<tr>
<td>14</td>
<td>1470</td>
<td>1841</td>
</tr>
<tr>
<td>15</td>
<td>1783</td>
<td>3599</td>
</tr>
<tr>
<td>16</td>
<td>1503</td>
<td>2401</td>
</tr>
<tr>
<td>17</td>
<td>1408</td>
<td>1728</td>
</tr>
<tr>
<td>18</td>
<td>1952</td>
<td>4782</td>
</tr>
<tr>
<td>19</td>
<td>1963</td>
<td>4506</td>
</tr>
<tr>
<td>20</td>
<td>2086</td>
<td>4273</td>
</tr>
<tr>
<td>21</td>
<td>1942</td>
<td>2265</td>
</tr>
<tr>
<td>22</td>
<td>3625</td>
<td>3653</td>
</tr>
<tr>
<td>23</td>
<td>1903</td>
<td>4944</td>
</tr>
<tr>
<td>24</td>
<td>1304</td>
<td>4039</td>
</tr>
<tr>
<td>25</td>
<td>1722</td>
<td>2679</td>
</tr>
<tr>
<td>26</td>
<td>1601</td>
<td>3542</td>
</tr>
<tr>
<td>27</td>
<td>2374</td>
<td>5517</td>
</tr>
<tr>
<td>28</td>
<td>1918</td>
<td>2163</td>
</tr>
<tr>
<td>29</td>
<td>1588</td>
<td>4657</td>
</tr>
<tr>
<td>30</td>
<td>1687</td>
<td>3972</td>
</tr>
</tbody>
</table>

Note. All scores represent the more complex (SO) condition of the task. The critical region for mathematical operations is Region 3; the critical regions for the visuo-spatial task are Regions 1 and 2. Aggregate scores for mathematical and visuo-spatial operations were obtained by averaging reaction times across all regions of each task.