

**A New Tool for Measuring and Understanding Individual Differences in the
Component Processes of Reading Comprehension**

by

Brenda Hannon

**A thesis submitted in conformity with the requirements
for the Degree of Doctor of Philosophy
Graduate Department of Psychology
University of Toronto**

© Copyright by Brenda Hannon (2001)



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*

Our file *Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-63776-X

Canada

A New Tool for Measuring and Understanding Individual Differences in the
Component Processes of Reading Comprehension

Doctor of Philosophy 2001

Brenda Hannon

Department of Psychology, University of Toronto

Abstract

This thesis proposes a new measure of individual differences in reading comprehension ability that is theoretically motivated, easy to administer, and has high predictive power. Participants read three-sentence paragraphs that describe the relations among a set of real and artificial terms such as: A NORT resembles a JET but is faster and weighs more., A BERL resembles a CAR but is slower and weighs more., A SAMP resembles a BERL but is slower and weighs more. By using the relations described in the paragraph, participants can construct linear orderings (e.g., for speed linear ordering NORT > JET > CAR > BERL > SAMP); however, because the fact that a JET is faster than a car is not explicitly mentioned, participants need to access their existing world knowledge to construct this ordering. After studying the paragraph, participants respond to true-false statements that assess their abilities on four separate component processes of reading comprehension. Text memory statements (e.g., A NORT is faster than a JET.) test memory for information explicitly mentioned in the paragraph; no prior knowledge is required. Text inferencing statements (e.g., A SAMP is slower than a CAR.) test implicit information that can be inferred by combining information that appears explicitly in the text (i.e., A BERL is slower than a CAR; A SAMP is slower than a BERL.) Knowledge access

statements (e.g., A JET is faster than a CAR.) can be answered by accessing prior knowledge; no new text information is required. Knowledge integration statements (e.g., A NORT weighs more than a CAR.) test implicit information that require participants to access their prior knowledge (i.e., a JET weighs more than a CAR) and integrate this knowledge with the text fact (i.e., A NORT weighs more than a JET). The components of the task predict performance on a test of global reading comprehension, and on a range of specific comprehension tests, each of which draws more heavily on one particular component. The component processes task is better at predicting reading comprehension than is a typical working memory span task, and has the potential of advancing the understanding and measurement of a range of linguistic and cognitive tasks.

Acknowledgments

I am indebted to Meredyth Daneman, whose expertise, encouragement, and support have been invaluable. I am also grateful to Eyal Reingold and Gus Craik for their advice along the way. I thank Candice Moore, Jayme Pickett, and Angela Kingston for their help with experiment preparation, Marilyne Ziegler for her programming advice, and Bruce Schneider for his statistical advice. Thanks also go to many individuals for their constant encouragement and friendship; in particular, Amy Siegenthaler, Bruce Oddson, Tonya Bergeson, France Archambault, Calvin Embrey, Jayne Carey, and Margaret McKinnon.

Table of Contents

Abstract	ii
Acknowledgments	iv
Table of Contents	v
List of Tables	x
List of Figures	xii
List of Appendices	xiii
Chapter 1: General Introduction	1
Psychometric Approach	1
Cognitive-Correlates Approach	4
Cognitive-Contents Approach	7
Cognitive-Training Approach	8
Cognitive-Components Approach	9
Approach used in my Research	11
Chapter 2: Introduction to Experiments that Develop and Validate the New	
Component Processes Task	13
The Background	13
Single-component Approaches to Understanding Reading Ability	15
Multicomponent Approaches to Understanding Reading Ability	17
The New Component Processes Task	19
Chapter 3: Experiment 1	21
Introduction	21

Rationale for Experiment	21
Description of the Potts and Peterson Task	21
Methods	23
Participants	23
Potts and Peterson Task	24
Materials	24
Procedure	26
Reading Comprehension Test	27
Vocabulary Knowledge Test	27
Verbal Analogies Test	27
Deductive Reasoning Test	28
General Analytic Reasoning Test	28
Results and Discussion	28
Correlations Among Components of Potts and Peterson Task	31
Regression Analyses on Reading Comprehension Test Scores using the Components of the Potts and Peterson Task as Predictors	32
Chapter 4: Experiment 2	35
Introduction	35
Design of Component Processes Task	35
Method	38
Participants	38
Component Processes Task	38

Materials	38
Procedure	42
Spatial Rotation Ability Test	43
Results and Discussion	43
Correlations Among Components of Component Processes Task	45
Reliability Estimates for Components of Component Processes Task ..	47
Correlations Between Components and Reading Comprehension	47
Regression Analyses on Reading Comprehension Test Scores using the Components of the Component Processes Task as Predictors	50
Regression Analyses of Reading Comprehension Scores with Vocabulary Knowledge and Component Processes as Predictors	54
Chapter 5: Experiment 3	56
Introduction	56
Rationale for Examining Predictive Power of Component Processes Task	56
Rationale for Examining Construct Validity of each Component	56
Method	58
Participants	58
Short-Stories Task	58
Verification Task	60
Bridging Inference Task	61
Inferring-New-Word-Meanings-From-Context Task	63

Memory-for-Details Task	64
Results and Discussion	65
General Results	66
Correlations Among Components of Component Processes Task .	66
Regression Analyses of Reading Comprehension Scores using	
Component Processes as Predictors	69
Short-Stories Task	70
Verification Task	76
Bridging Inference Task	76
Inferring-New-Word-Meanings-From-Context Task	78
Memory-for-Details Task	79
Summary of Results	80
Chapter 6: Experiment 4	81
Introduction	81
Rationale for Comparing Predictive Power of Working Memory Span	
Test with Components of Component Processes Task	81
Description of Daneman and Carpenter Working Memory Span Task	81
Methods	83
Participants	83
Working Memory Span Test	83
Results and Discussion	84
Correlations among Components of Component Processes Task	84

Regression Analyses of Components and Reading Comprehension	84
Comparing Predictive Power of Working Memory Span Task and Component Processes Task	88
Summary of Results	89
Chapter 7: Analysis of Pooled Data from the Previous Experiments	90
Correlations Among Components of Component Processes Task	90
Confirmatory Factor Analyses	92
Rationale for Confirmatory Factor Analysis	92
Measurement Model	92
Structural Equation Model	94
Regression Analyses	96
Chapter 8: General Summary and Conclusions	98
References	103
Appendix A: Practice Paragraph, Experimental Paragraphs, and Test Statements for Component Processes Task	116
Appendix B: Instructions for the New Component Processes Task.....	132

List of Tables

Table		Page
1	Means and Standard Deviations for Tasks used in Experiment 1	30
2	Correlations Among the Four Components in the Potts and Peterson Task used in Experiment 1	31
3	Correlations between the Potts and Peterson Components and Other Verbal Abilities Tests in Experiment 1	32
4	Means and Standard Deviations for Tasks used in Experiment 2	44
5	Correlations Among the Components of the Component Processes Task in Experiment 2	46
6	Correlations between Components of the Component Processes Task and Abilities Tests in Experiment 2	49
7	Regression Analyses on Reading Comprehension Test Scores using the Components of the Component Processes Task as Predictors	53
8	Means and Standard Deviations for Tasks in Experiment 3	67
9	Correlations among the Components of the Component Processes Task and Reading Comprehension in Experiment 3	68
10	Correlations Among the Components of the Component Processes Task and Specific Reading Comprehension Tasks in Experiment 3	71
11	Regression Analyses on Specific Reading Comprehension Tasks with Components of Component Processes Task as Predictors	72
12	Means and Standard Deviations for Tasks used in Experiment 4	86

Table		Page
13	Correlations among Components of Component Processes Task, Reading Comprehension, and Working Memory Span in Experiment 4	87
14	Regression Analyses on Reading Comprehension Scores in Experiment 4 with Working Memory Span and Component Processes as Predictors	89
15	Correlations Among the Components of the Component Processes Task and Reading Comprehension across Experiments 2,3, and 4	91
16	Results of CFA Analyses for Measurement and Structural Models	93
17	Regression Analyses on Reading Comprehension Scores for all 222 Participants with Components of Component Process Task as Predictors	97

List of Figures

Figure		Page
1	Path Diagram Showing the Measurement Model and the Structural Equation Model for the Component Processes Task (n=222).	95

List of Appendices

Appendix A: Practice Paragraph, Experimental Paragraphs, and Test Statements for Component Processes Task	117
Appendix B: Instructions for the New Component Processes Task	132

Chapter 1: General Introduction

What is reading comprehension? What accounts for individual differences in reading comprehension ability? And how might individual differences in reading comprehension ability be measured? In this thesis, I argue that there are four main component processes underlying reading comprehension ability and I propose a new tool that measures individual differences in these four theoretically motivated component processes. The purpose of this chapter is to discuss the general approaches that have been used to study individual differences in cognitive abilities such as reading comprehension ability, and to show how my research approach fits into these general approaches. These five approaches are: (a) the psychometric approach, (b) the cognitive-correlates approach, (c) the cognitive-contents approach, (d) the cognitive-training approach, and (e) the cognitive components approach (see Sternberg, 1979). The psychometric approach focuses on the measurement of individual differences in cognitive abilities such as general intelligence, verbal abilities, and spatial abilities. The four other approaches apply information-processing theory towards understanding the nature of these individual differences.

The Psychometric Approach

Investigators using a psychometric approach to understand individual differences in cognitive abilities are interested in understanding cognitive abilities in terms of a set of static latent sources of individual differences called factors (see Sternberg, 1985, for review). The idea here is that individual differences in a cognitive ability (e.g., intelligence) can be decomposed into individual differences in factors, each of which is proposed to represent a distinct human ability (e.g., verbal ability, mathematical ability, spatial ability).

Factors are hypothetical constructs that are intended to explain the underlying sources of individual differences that give rise to the observed individual differences in test scores. To obtain factors, a technique called factor analysis is used; it involves calculating correlations among scores on tests, and then reducing these correlations to just a few linear combinations called factors. For example, if tests measuring vocabulary, reading comprehension, arithmetical problem solving, and arithmetical concepts were factor analyzed, the correlations among the scores on these tests would probably be reduced to two factors: verbal ability and mathematical ability (see Sternberg, 1985).

Theorists who have used the psychometric approach to understand individual differences in intelligence differ greatly in the number of factors they propose. At the lower end, Spearman (1927) proposed a single factor called "g", which he believed to permeate performance on all intellectual tasks. Thurstone (1938) proposed a more "middle-of-the-road" position by positing that intelligence is comprised of seven "primary mental abilities" (i.e., verbal comprehension, verbal fluency, numbers, spatial visualization, memory, reasoning, and perceptual speed). Finally, Guilford (1981) proposed the upper extreme position by positing that intelligence is comprised of 150 distinct factors. According to Guilford, every mental task involves three categories: (a) an operation, (b) a content, and (c) a product. Since there are five operations (cognition, memory, divergent production, convergent production, and evaluation), five kinds of contents (visual, auditory, symbolic, semantic, and behavioural), and six kinds of products (units, classes, relations, systems, transformations, and implications) and each of the subcategories are independently defined, they are multiplicative; so there are $5 \times 5 \times 6 = 150$ different mental abilities.

During the 1930's and 1940's, the psychometric approach was also used to develop a number of tests of reading comprehension ability, and like the research identifying the factors underlying intelligence, a rather lively debate ensued about what the sources of individual differences in reading comprehension might be. For example, Hunt (1947) and Robinson and McCollom (1934) proposed two-factor explanations of reading comprehension ability; for Hunt (1947) the two factors were (a) vocabulary and (b) a common reading comprehension factor, whereas for Robinson and McCollom (1934) the two factors were (a) rate of reading and (b) accuracy of comprehension. At the upper extreme was Davis (1944), who proposed a nine-factor explanation of reading comprehension ability. His nine factors were: (a) knowledge of word meanings, (b) the ability to select the appropriate meaning for a word in light of its particular contextual setting, (c) the ability to follow the organization of a passage and to identify antecedents and references in it, (d) the ability to select the main thought of a passage, (e) the ability to answer questions that are specifically answered in a passage, (f) the ability to answer questions that are answered in a passage but not in the words in which the question is asked, (g) the ability to draw inferences from a passage, (h) the ability to recognize the literary devices used in a passage and to determine its tone and mood, and finally, (i) the ability to determine a writer's purpose, intent, and point of view.

The psychometric approach has made great strides in measuring and statistically demonstrating that there are a number of potential sources of individual differences in reading comprehension ability, but the earlier applications of this approach were not without their problems. One problem with the earlier applications of the psychometric approach was the lack of any theoretical framework for understanding the cognitive constructs under investigation.

Instead, researchers relied on their own judgements for interpreting the factors underlying the cognitive construct (see Davis, 1944 for an example of this problem). In fact, theory and measurement were not integrated until the mid 1970's when researchers started applying information-processing theory toward identifying individual differences in cognitive abilities. There are four information-processing approaches to understanding individual differences in cognitive abilities: (a) the cognitive-correlates, (b) the cognitive-training, (c) the cognitive-contents, (d) and the cognitive components approaches. Each approach attempts to capitalize on individual differences in cognitive abilities in terms of the architecture and processes of the human information-processing system.

The Cognitive-Correlates Approach

Investigators using a cognitive-correlates approach to understand individual differences in cognitive abilities are interested in determining the extent to which tasks that contemporary information-processing psychologists believe measure basic human information-processing abilities correlate with performances on complex cognitive tasks. Researchers have investigated whether performance on basic information-processing tasks such as Posner and Mitchell's (1967) name-matching task, Sternberg's (1966) memory-scanning task, and Peterson and Peterson's (1959) short-term memory task correlate with performance on tests of verbal abilities and reading comprehension ability such as the Washington Pre-College Test (e.g., Hunt, Lunneborg, & Lewis, 1975) and the Nelson-Denny test of reading comprehension ability (e.g., Palmer, MacLeod, Hunt, & Davidson, 1985). For example, the Posner and Mitchell task involves the presentation of two letters identical in both name and physical type (A, A), identical in name but not case (A, a), or different in both name and case (A, b). Participants are presented letter pairs

and their reaction times to answer whether the two letters are the same or different are measured. On name identify (NI) trials, they are asked to respond "same" if the two letters share the same name and on physical identity (PI) trials they must respond "same" only if the two letters are physically identical. Of major interest is the difference between the time taken to respond "same" when the letters are identical in name only (e.g., A, a) versus when they are identical in both name and physical form (A, A). This NI-PI difference, typically 50-100 msec, is assumed to indicate the additional time required in the name identity situation for accessing the name of the visual letter code in memory. The time required to access the name of a letter appears to differ among individuals. If individuals who perform better on the complex tests of verbal abilities also show a smaller NI-PI difference, this would suggest that they are faster at retrieving well-learned letter or name codes from long-term memory.

The cognitive-correlates approach has been successful in demonstrating that theoretically motivated basic information-processing tasks are correlated with tasks assumed to measure verbal ability and reading comprehension ability. For example, using the basic Posner and Mitchell name-matching task, researchers have demonstrated that the NI-PI difference effect is smaller for high verbal individuals, suggesting that high verbal individuals are faster at retrieving well-learned letter or name codes from memory (Hunt et al., 1975). High verbal participants are faster in other retrieval processes as well. They are faster in other semantic matching tasks, such as responding "same" if two words are synonyms, or if two words are homonyms (Jackson & McClelland, 1979). However, Jackson and McClelland found that better readers were not faster at matching dot patterns, suggesting that the processing speed advantage of better readers did not lie in the encoding or comparing of nonverbal visual displays.

Despite these encouraging findings, the initial applications of the cognitive-correlates approach toward understanding individual differences in verbal abilities have been heavily criticized. Specifically, critics have questioned whether the theories underlying the basic information-processing tasks are the best theoretical framework for interpreting the complex information processing required by many psychometric tests (Sternberg, 1981). Typically, the correlations between the basic information-processing tasks and the tasks measuring complex abilities are rarely above .30 (i.e., rarely capture more than 9% of the variance), and critics have argued that these relatively low correlations occur because tests measuring complex cognitive abilities draw on these lower level perceptual and memory abilities in only a peripheral way (Sternberg, 1981).

More recent applications of the cognitive-correlates approach have avoided this criticism by replacing the basic information-processing tasks with more complex information-processing tasks that are designed to tap more sophisticated cognitive processes and resources. For example, Daneman and Carpenter (1980) showed that their more complex working memory span task provided a more suitable theoretical framework for understanding reading comprehension ability than the basic word span task. The working memory span task requires participants to simultaneously coordinate both the processing and storage requirements that are imposed. For example, the processing requirements might be to read aloud sets of unrelated sentences (e.g., An eerie breeze suddenly chilled the warm humid air. The umbrella grabbed its bat and stepped up to the plate. His mouth was twisted into an inhuman smile.) and to make judgments about the sensibility of each sentence, and the storage requirements might be to remember the sentence final words (e.g., air, plate, smile) so that they can be recalled at the end of the set. On the other

hand, the basic word span task imposes only a storage requirement by having participants store and recall back a random string of words (e.g., cup, shoe, ball). Daneman and Carpenter (1980) showed that performance on the working memory span task was a much better predictor of performance on the Verbal Scholastic Aptitude Test (VSAT) than the basic word span task. Indeed, Daneman and Carpenter (1980) found that their working memory span tasks accounted for 56% of the variance on standard tests of reading comprehension ability (see Daneman & Merikle, 1996).

The Cognitive-Contents Approach

Investigators using a cognitive-contents approach to investigate individual differences in cognitive abilities are interested in comparing the performances of experts and novices on complex tasks. The idea is that experts and novices perform differently on complex tasks because they differ in both the contents and structures of their knowledge, and if these differences are identified, then important sources of individual differences might be revealed.

The cognitive-contents approach has been used to study expert-novice differences in the selection of moves and strategies in chess (e.g., Chase & Simon, 1973) and memory for digits (e.g., Chase & Ericsson, 1982). This approach has also been used by researchers in the area of verbal abilities. For example, Spilich, Vesonder, Chiesi, and Voss (1979) compared performances between participants with high domain knowledge (e.g., experts) and low domain knowledge (e.g., novices) on a task that measured memory for a text about a baseball game. They observed that experts recalled larger chunks of information for baseball actions and had more accurate accounts of the order of game actions that were related to the goal structure than did novices. Spilich et al. (1979) interpreted these findings as evidence that previously acquired

knowledge affects the processing of new domain-related information (see Korkel & Weinert, 1989 and Yekovich, Walker, Ogle, & Thompson, 1990 for other applications of the cognitive-contents approach). A major difference between the cognitive-contents approach and the cognitive-correlates approach is that the former emphasizes content and structure rather than cognitive processes.

The Cognitive-Training Approach

Investigators using a cognitive-training approach to investigate individual differences in cognitive abilities examine how much the training on a single cognitive component improves overall task performance. The idea is that cognitive training of a single component might improve overall task performance, and if overall task performance does improve then the trained component is an important source of individual differences.

The cognitive-training approach has had mixed success at enhancing verbal abilities. For example, although most vocabulary-building training studies are successful at increasing vocabulary knowledge, there is little concomitant increase in reading comprehension ability, despite the fact that vocabulary knowledge is considered to be one of the best single predictors of reading comprehension ability (see Beck, Perfetti, & McKeown, 1982 for a summary of the literature). Rather, it seems that for vocabulary training to enhance reading comprehension ability, it must include rigorous training that focuses on increasing processing fluency (Beck et al., 1982). Similarly, Pressley and Brewster (1990) demonstrated that an elaboration strategy called elaborative interrogation increases knowledge of a domain (see also Woloshyn, Willoughby, Wood, & Pressley, 1990; and Woloshyn, Pressley, & Schneider, 1992) when participants are asked "why" questions about that knowledge domain. However, in order to

ensure continued use of elaborative interrogation, training must include information about the strategy's effectiveness and metacognitive information about why the strategy works (Pressley, 1994).

At a practical level, the cognitive-training approach is useful for revealing what aspects of training (e.g., feedback on success of strategy, metaknowledge about the strategy) are most effective for improving an individual's performance. However, with the exception of elaborative interrogation (e.g., Pressley & Brewster, 1990) which I interpret as a cognitive training intervention for knowledge-based inferencing (see also Hannon & Daneman, 1998), the cognitive-training approach fails to generate new theories about cognitive abilities. Rather it seems that its real value is when it is used in conjunction with one of the other approaches. Such a combination of approaches involves establishing a theory for a cognitive ability using either the cognitive-correlates, cognitive-contents, or cognitive-components approach, developing training methods that enhance one or more cognitive components of the pre-established theory, and then demonstrating that training on the cognitive components causes measurable changes in overall task performance.

The Cognitive-Components Approach

Investigators using a cognitive-components approach for investigating individual differences in cognitive abilities are interested in isolating component processes of complex subtasks (e.g., analogies) that are important for overall performance on a psychometric measure (e.g., intelligence). The major focus of this approach is not simply to show that performance on each subtask correlates with performance on the psychometric measure of the cognitive ability. Rather, the idea is that there are a few important component processes that are common among

the subtasks of a psychometric measure, and if these component processes and the relationships among these component processes are identified, then a theoretical framework for the cognitive ability of interest can be developed. For example, Sternberg (1977) measured the processing times required at different stages for the solutions of analogies such as LAWYER : CLIENT :: DOCTOR : (a) PATIENT, (b) MEDICINE. Using a subtractive method to isolate the processing times for the component processes of (a) encoding and formulating a general strategy, (b) inferring the relations among some of the terms in the analogy, (c) mapping the inferred relations to solution choices, and (d) applying the previously inferred relations to the solution choices, Sternberg observed that the component process for encoding predicted performance on other traditional tests of intelligence (Sternberg 1977; Sternberg, 1996). Specifically, he observed that the slower participants were at the component process of encoding on the analogies test, the higher they scored on other traditional tests of intelligence (Sternberg, 1977; Sternberg, 1996).

The cognitive-components approach has also been used for identifying the component processes of reading comprehension ability in adolescents (e.g., Frederiksen, 1982). Frederiksen (1982) selected a number of reading subtasks that he classified as measures of word analysis, context utilization, and discourse analysis. Then based on the theory of the reading subtasks, Frederiksen decomposed each subtask into component processes. For example, one of the word analysis subtasks called pseudoword pronunciation, was decomposed into two component processes: perceiving multi-letter units and decoding or phonological translation in word recognition. After examining the relationships among the component processes of the subtasks, Frederiksen examined the extent that each component process predicted performance on a test of reading comprehension ability. Frederiksen (1982) observed that for adolescents,

decoding/phonological efficiency and activation of semantically-related concepts in memory were the best predictors of performance on the test of reading comprehension ability.

The cognitive-components approach has received some harsh criticism with respect to tendencies to formulate theories based on experimentally-defined, systematic breakdowns of the subtasks. Indeed Neisser (1983) has argued that the theories are just solution steps: "By now it should be clear that the products of Sternberg's method are not 'components' at all, but solution steps that emerge in response to the demands of special experimental procedures. It should be equally clear that the componential theory is not a theory of intelligence but a method of running subjects and analyzing data that yields stepwise solutions." (pp.196) While this criticism may apply to Sternberg's use of componential analysis of analogies, it is less applicable to Frederiksen's (1982) use of componential analysis. In Frederiksen's research, the components of the reading subtasks were identified using pre-established theoretical interpretations and not solution steps for executing the tasks.

Approach used in my Research

The approach that I used for measuring and understanding individual differences in reading comprehension ability was componential in nature. Like previous applications of the componential approach, the focus of my research was to use the theory of the component processes to inform me about the nature of the psychological construct (e.g., reading comprehension ability). In addition, the methodology I used was similar to Sternberg's (1977) and Frederiksen's (1982) inasmuch as the important component processes of a larger psychological construct (e.g., reading comprehension) were identified, the relationships among the components were examined, and then the independent and overlapping contributions of the

component processes to overall performance on the psychometric measure (e.g., reading comprehension) were examined.

On the other hand, my application of componential analysis is more similar to Frederiksen's application of the componential approach than it is to Sternberg's application of the componential approach. Like Frederiksen, the component processes in my new measure are theoretically motivated and not experimentally defined. This allowed me to avoid Neisser's criticism of componential analyses as simply a method of devising experimentally defined, solution steps. However, whereas Frederiksen (1982) broke reading subtasks down into lower level component processes such as letter encoding efficiency, decoding or phonological translation, the component processes used in my new measure are higher level component processes (text memory, text inferencing, knowledge access, and knowledge-based inferencing). The advantage of my approach is that it allowed me to examine, for the first time, the contributions of higher level component processes to a psychological construct such as reading comprehension. Finally, unlike Frederiksen (1982) who used separate subtasks that were believed to measure some of the important component processes of the construct under investigation, my approach used a single task that measures the four component processes of the psychological construct under investigation. The advantage of my approach is that it avoids giving multiple instructions for multiple tasks. Rather, only one set of instructions are provided, making the new measure practical and easy to administer.

Chapter 2: Introduction to Experiments

In this thesis I propose a new individual-differences tool that provides estimates of a reader's ability to access and integrate long-term memory knowledge with text information, to make text-based inferences, and to recall text. This tool is based on a rather simple but ingenious task developed by Potts and Peterson (1985). The modified version of the task is easy to administer and accounts for a substantial proportion of the variance in performance of university students on a global, standardized test of reading comprehension ability. First, I briefly describe some of the previous approaches that have been used to measure and understand reading comprehension ability, and I highlight the advantages of the new task. Then, I describe four studies that were conducted as part of the process of developing and validating the component processes task.

The Background

Research on reading comprehension has shadowed research on intelligence. This is not surprising, given that reading comprehension is highly correlated with general intelligence and both are related to school success (Bloom, 1976; Perfetti, 1976). One can trace in the histories of both research movements a similar tension between concerns with measurement on the one hand and theory on the other. Investigations of reading, like intelligence, had their origin primarily in the practical concerns of measurement and prediction. Researchers were interested in quantifying mental abilities in order to predict performance in school, business and industry, government and the military. This interest gave birth to the mental testing movement and a plethora of standardized tests, not only to assess intelligence, but reading comprehension as

well—tests like the Davis Reading Test, the Metropolitan Reading Test, the Nelson-Denny, and the Verbal Scholastic Aptitude Test (VSAT), to name but a few. Many of the tests predicted performance with considerable accuracy and reliability. However, one consequence of the testing movement's goal of pursuing predictive validity was that construct validity was lost in the shuffle. Because reading comprehension tests were largely atheoretical, there was no consensus on what was being measured. Nevertheless, this did not deter researchers from persisting with their business of measuring. Presumably, they were confident that the tests had to be measuring something important, because many of the tests were undeniably reliable and predictive.

With the advent and dominance of the information-processing approach to cognition, the emphasis has switched from measurement to theory. The goal is no longer simply to quantify individual differences in intellectual tasks, but also to explain individual differences in terms of the architecture and processes of the human information-processing system (Daneman, 1982; Sternberg, 1985). In the field of intelligence, models have been developed to account for the basic processes and components underlying performance on psychometric tests (e.g., Carroll, 1976, 1978; Hunt, 1980; Kyllonen, 1993; Kyllonen & Christal, 1990; Sternberg, 1979, 1980, 1985). Similarly, in the field of reading and language comprehension, there have been attempts to account for the processes and components that might differentiate skilled from less-skilled readers (e.g., Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Carr, 1981; Chiang & Atkinson, 1976; Hunt, 1978; Jackson & McClelland, 1979; Palmer, MacLeod, Hunt, & Davidson, 1985; Perfetti & Lesgold, 1977; see Daneman, 1991 for a review). Because reading is a complex cognitive skill that draws on many component processes and resources, any of these component processes or resources has the potential for being a source of individual differences in

reading ability. Some theories of reading ability have emphasized a single component as the major source of individual differences in reading ability. Other theories have emphasized a more multicomponent approach (see Carr, 1981).

Single-Component Approaches to Understanding Reading Ability

Most theories of reading ability have promoted a single component as the major source of individual differences in reading comprehension performance. However, there has been little agreement on what that component is. The evidence suggests that skilled and unskilled readers do not differ in the extent of their control over how they move and place their eyes during reading (Just & Carpenter, 1987; Olson, Kliegl, & Davidson, 1983; Rayner & Duffy, 1988; Stanley, Smith, & Howell, 1983; although see Pavlidis, 1981), nor do they differ in the amount of visual information they can extract from a single fixation (Samuels, LaBerge, & Bremer, 1978; Underwood & Zola, 1986). However, they do differ on many of the other component processes of reading.

Some theories emphasize word knowledge or word recognition skills as the major source of individual differences in reading ability. It has been argued that less-skilled readers have smaller vocabularies than skilled readers (Davis, 1968; Thorndike, 1973), are slower and less efficient at recognizing written words (Perfetti, 1985; Stanovich, Cunningham, & Feeman, 1984), are slower at accessing word meanings from semantic memory (Hunt et al., 1975), and are poorer at deriving phonology from print (Hogaboam & Perfetti, 1978; Frederiksen, 1982; Snowling, 1980).

Other theories of reading ability emphasize the higher level language comprehension processes that compute the semantic, syntactic, and referential relationships among successive

words, phrases, and sentences in a text. It has been argued that less-skilled readers are at a particular disadvantage when they are required to execute a process that requires integrating newly encountered information with information encountered earlier in the text or retrieved from long-term memory (Daneman, 1991). So for example, less-skilled readers have problems interrelating successive topics (Lorch, Lorch, & Morgan, 1987) and integrating information to derive the overall gist or main theme of a passage (Palincsar & Brown, 1984). They have more difficulty computing the antecedent referent for a pronoun (Oakhill & Yuill, 1986). They have more difficulty making inferences, and they tend to make fewer of them spontaneously during reading (Hannon & Daneman, 1998; Long, Oppy, & Seely, 1994; Oakhill, 1982). Indeed, less-skilled readers tend not to demand informational coherence and consistency in a text, and they often fail to detect, let alone repair, semantic inconsistencies (Garner, 1980).

Whereas these higher level theories of reading ability have focused on a specific integration process of reading, other higher level theories have focused on a resource shared by many of the integration processes. According to working memory theories of reading ability, less-skilled readers are at a disadvantage at all of the processes that require the integration of newly encountered information with information encountered earlier in the text because they have less capacity to keep the earlier information active in temporary storage (Daneman & Carpenter 1980, 1983; Engle, Cantor, & Carullo, 1992; Just & Carpenter, 1992; see Daneman & Merikle, 1996, for a review) . According to knowledge-based theories of reading ability, less-skilled readers are at a disadvantage at all of the processes that require the integration of newly encountered information with information retrieved from long-term memory (Anderson & Pearson, 1984; Spilich et al., 1979; Spiro, 1980; Voss, Fincher-Kiefer, Greene, & Post, 1985).

The single-component research has successfully identified a whole range of cognitive and linguistic components and resources that are correlated with reading comprehension ability. However, the single-component approach does not allow researchers to determine the relative contributions of individual differences in the various components to individual differences in overall reading comprehension ability. Nor does the single-component approach allow researchers to determine the extent to which the various components make independent contributions to reading comprehension ability.

Multicomponent Approaches to Understanding Reading Ability

Recognizing that no single component can account for all the variance in performance on a complex task such as reading, some researchers have administered a battery of perceptual, cognitive, and linguistic tasks to their readers in an attempt to determine the relative amounts of variance in reading ability uniquely associated with each component process (e.g., Baddeley et al., 1985; Bell & Perfetti, 1994; Carr, 1981; Dixon, LeFevre, & Twilley, 1988; Engle, Nations, & Cantor, 1990; Haenggi & Perfetti, 1994; Jackson & McClelland, 1979; Palmer et al., 1985; Rankin, 1993). One fairly general finding is that the lower level word recognition and lexical access processes account for relatively little of the variance in reading comprehension performance once readers get beyond the beginning stages of reading (Chabot et al., 1984; Jackson & McClelland, 1979; Palmer et al., 1985). Rather, it is the higher level processes common to both reading and listening comprehension that account for a substantial proportion of the variance in reading comprehension ability (Jackson & McClelland, 1979; Palmer et al.,

1985).¹ Also, the studies have shown that some component processes make independent contributions to aspects of reading comprehension ability. For example, Dixon et al. (1988) showed that vocabulary knowledge and working memory capacity make independent contributions to reading comprehension performance. And Haenggi and Perfetti (1994) showed that prior knowledge is related to answering explicit questions about expository text, whereas a probe discourse working memory task is related to answering questions of a more implicit nature.

Multicomponent research has advanced the understanding of the nature of individual differences in reading ability by accounting for considerable proportions of variance in performance on the atheoretical standardized tests of reading comprehension ability, and by providing information concerning the relative amounts of variance in reading ability uniquely associated with each of the component processes and resources. Nevertheless, I believe that the tension between theory and measurement has now shifted too much in the direction of theory. The multicomponent approach has provided a better understanding of the kinds of cognitive skills and processes being tapped by global, standardized test of reading comprehension ability such as the Nelson-Denny and the VSAT; however, it has come at the expense of developing new measures of reading comprehension ability that are theoretically motivated and practical to administer. So if researchers wish to obtain an accurate assessment of someone's reading ability, they must chose between administering a large battery of measures that taps theoretically important component processes and resources of reading but is cumbersome and time consuming

¹Several studies have shown that the speed of recognizing words and accessing meanings tends to be more closely related to measures of reading speed than to measures of reading comprehension (Chabot et al., 1984; Jackson & McClelland, 1979; Palmer et al., 1985).

to administer, or they must rely on the standardized but atheoretical tests of reading comprehension ability.

The New Component Processes Task

My new task has the advantages of being both theoretically motivated and practical to administer. The task measures individual differences in four components of reading comprehension—the ability to access prior knowledge from long-term memory, to integrate accessed prior knowledge with new text information, to make inferences based on information provided in the text, and to recall the new text information from memory. Of course, the task is quicker and easier to administer than a battery of tasks measuring many components. Nevertheless, it accounts for a substantial proportion of the variance on the Nelson-Denny standardized test of reading comprehension ability. Thus, it has both construct validity and predictive power.

In Experiment 1 (Chapter 3), I demonstrate that an existing four-component task developed by Potts and Peterson (1985) has only low to moderate correlations with reading comprehension ability and other tests of verbal abilities. In Experiment 2 (Chapter 4), I introduce the modified version of the Potts and Peterson (1985) task and show that it is a much better predictor of global reading comprehension performance than is the original Potts and Peterson (1985) version. In Experiment 3 (Chapter 5), I validate each component of the task by showing that the text memory component is the best predictor of performance on a memory-loaded reading task, the text inferencing component is the best predictor of performance on an inference-loaded reading task, and so on. In Experiment 4 (Chapter 6), I compare the predictive power of the component processes task with that of working memory span, another theoretically

motivated measure that is a good predictor of reading comprehension ability. In Chapter 7, I pool together and analyze the data from all the participants who were administered the task (Experiments 2, 3, and 4 combined), and I present the results of a confirmatory factor analysis for the structure of the component processes task. Finally in Chapter 8, I consider the potential applications of the task toward advancing the understanding and measurement of reading comprehension ability.

Chapter 3: Experiment 1

Experiment 1 tested whether a task developed by Potts and Peterson (1985) is related to reading comprehension ability. Potts and Peterson's (1985) task was originally developed for purposes other than predicting reading ability (see also Potts, 1977). However, on the basis of the findings of previous research on reading ability, I was struck by the possibility that all four components of Potts and Peterson's task might tap important sources of individual differences in reading comprehension ability. For example, Masson and Miller (1983) observed that both text memory and text inferencing are two important sources of individual differences in reading comprehension ability, Chabot et al. (1984) observed that skill at accessing prior knowledge is an important source of individual differences in reading comprehension ability, and Singer and Richot (1996) and Hannon and Daneman (1998) observed that integrating prior knowledge into text is an important source of individual differences in reading comprehension ability. Consequently, I administered the Potts and Peterson (1985) task to university-level students and examined the extent to which each of the four components was correlated with overall reading comprehension ability. I first briefly describe the Potts and Peterson (1985) task, and then I describe my study.

In the Potts and Peterson (1985) task, participants read short paragraphs, each consisting of three sentences that describe the relations among a set of real and artificial terms, such as: A JAL is larger than a TOC., A TOC is larger than a PONY., A BEAVER is larger than a CAZ. By using the relations described in the three sentences, participants can construct a five-item linear ordering (JAL > TOC > PONY > BEAVER > CAZ); however, because the fact that a pony

is larger than a beaver is not explicitly mentioned, participants need to access their existing world knowledge in order to construct the ordering. Participants study the paragraph at their own pace and then respond to true-false statements of four types. Text memory statements (e.g., A JAL is larger than a TOC) test memory for information explicitly mentioned in the paragraph; no prior knowledge is required. Text inferencing statements (e.g., A JAL is larger than a PONY) test for information that can be inferred by integrating information that appears explicitly in the text (i.e., A JAL is larger than a TOC; a TOC is larger than a PONY); again, no prior knowledge is required. Knowledge access statements (e.g., A PONY is larger than a BEAVER) can be answered by simply accessing prior knowledge; no new information from the text is required. Knowledge integration statements (e.g., A TOC is larger than a BEAVER) require participants to access their prior knowledge (i.e., a pony is larger than a beaver) and integrate this knowledge with the text fact (i.e., A TOC is larger than a PONY).

Potts and Peterson (1985) found that the pattern of correlations among the four statement types was consistent with the above description. Text memory and text inferencing, the two components that depended on memory for new text information rather than access to prior knowledge, were highly correlated with one another ($r = .59$), but neither was correlated with knowledge access, the component that depended only on prior knowledge ($r = -.04$ for text memory and $r = .00$ for text inferencing). On the other hand, knowledge integration, the component that depended on text information as well as prior knowledge, was correlated with the two text-based components ($r = .39$ with text memory and $r = .44$ with text inferencing), as well

as the pure prior knowledge component, knowledge access ($r = .64$).² This pattern of correlations suggests that the ability to remember new information and the tendency to use world knowledge are separate skills, and I was curious about the extent to which each is related to overall reading comprehension ability.

So far, researchers have focused on Potts and Peterson's (1985) knowledge integration component. For example, Potts and Peterson (1985) showed that individuals who performed well on the knowledge integration component were more likely to incorporate new information into their existing world knowledge than were individuals who performed poorly on the knowledge integration component. And Singer and Ritchot (1996) showed that individuals who performed well on the knowledge integration component were better at constructing bridging inferences during reading (see also Singer, Andrusiak, Reisdorf, & Black, 1992). However, based on previous individual-differences studies of reading ability (e.g., Chabot et al., 1984; Haenggi & Perfetti, 1994; Hannon & Daneman, 1998; Masson & Miller, 1983; Oakhill, 1982; see Daneman, 1991, for a review), it was hypothesized that all four components of the task might capture aspects of reading comprehension ability, and so the Potts and Peterson task was administered in conjunction with a widely-used global test of reading comprehension, the Nelson-Denny, as well as several other tests of verbal abilities.

Method

Participants

The participants were 57 University of Toronto students. All participants were fluent

²The correlations cited here are from Experiment 1 in Potts and Peterson's (1985) study; a similar pattern of correlations was observed in their second experiment.

speakers of English and were tested individually in two sessions. Participants were administered the Potts and Peterson (1985) task followed by five tests that measured: (a) reading comprehension ability, (b) vocabulary knowledge, (c) verbal analogies, (d) deductive reasoning, and (e) general analytic reasoning ability.

Potts and Peterson Task

As in Potts and Peterson (1985), participants read three-sentence paragraphs such as “the pony” one, and then answered 18 true-false test statements about it.

Materials. The materials consisted of a set of ten short paragraphs. The first two were used as practice paragraphs. The remaining eight consisted of the four paragraphs from the original Potts and Peterson (1985) task, as well as four new paragraphs modelled after the Potts and Peterson ones. Each paragraph consisted of three sentences that were presented one under the other simultaneously on a computer screen. Each three-sentence paragraph included three nonsense terms, two real terms, and a relation among them. For example, in “the pony” paragraph, JAL, TOC, and CAZ are the three nonsense terms, PONY and BEAVER are the two real terms, and larger than is the relation. By using the larger than relation, a participant could construct the five-item linear ordering of JAL > TOC > PONY > BEAVER > CAZ. However, because the fact that a pony is larger than a beaver was never actually mentioned, participants had to use their existing world knowledge about the relative size of a pony and a beaver to construct the linear ordering. The known relations used in Potts and Peterson’s (1985) other three experimental passages were: a plane is faster than a car, a piano is heavier than a chair, and

a grapefruit is larger than an orange (see Potts & Peterson, 1985).³ The known relations used in the four new paragraphs were: a scream is louder than a whisper, a tiger is wilder than a kitten, a giraffe is taller than a zebra, and a rope is stronger than a string.

After studying a paragraph, participants responded to 18 true/false statements about it. Half of the statements were true and half were false. There were four types of test statements: text memory, text inferencing, knowledge access, and knowledge integration.

The text memory statements tested memory for information explicitly presented in the paragraph. No prior knowledge was required. For example, a true text memory statement for the “pony” paragraph was A TOC is larger than a PONY. Each true text memory statement had a corresponding false statement that was constructed by reversing the order of the terms in the true statement, such as A PONY is larger than a TOC. Each paragraph had three true and three false text memory statements.

The text inferencing statements required participants to make inferences about information presented explicitly in the paragraph. No prior knowledge was required. For example, the true text inferencing statement for the “pony” paragraph was A JAL is larger than a PONY. This statement required drawing the inference from information explicitly mentioned in the paragraph sentences, A JAL is larger than a TOC, and A TOC is larger than a PONY. Each true text inferencing statement had a corresponding false text inferencing statement, such as A PONY is larger than a JAL. Each paragraph had one true and one false text inferencing statement.

³Potts and Peterson (1985) verified that these facts were part of a college student’s world knowledge by testing an independent group of undergraduates for their knowledge of them.

The knowledge access statements tested access to prior knowledge. No information from the paragraph was required. For example, the true knowledge access statement from the “pony” paragraph was A PONY is larger than a BEAVER. This statement required participants to access their prior knowledge that a pony is larger than a beaver because this fact was not explicitly mentioned in the paragraph. Each true knowledge access statement had a corresponding false knowledge access statement, such as A BEAVER is larger than a PONY. Each paragraph had one true and one false knowledge access test statement.

Finally, the knowledge integration statements tested integration of prior knowledge with text information. For example, a true knowledge integration statement from the “pony” paragraph was A TOC is larger than a BEAVER. This statement required participants to access their prior knowledge that a pony is larger than a beaver and integrate this knowledge with the text information that A TOC is larger than a PONY. Each true knowledge integration statement had a corresponding false knowledge integration statement, such as A BEAVER is larger than a TOC. Each paragraph had four true and four false knowledge integration statements.

Procedure. As in Potts and Peterson (1985), participants were told that the primary interest of this study was to investigate how people use what they already know to help them learn new information. They were informed that when they pressed the space bar on the computer keyboard, a to-be-learned paragraph consisting of three sentences would be presented on the computer screen. They were told that some of the terms in the sentences would represent real items that should be familiar to them, and that other terms, denoted by nonsense syllables, would represent imaginary items unfamiliar to them. Then, after they had learned the paragraph, they were required to respond to a series of true–false test statements about the paragraph.

Participants were explicitly instructed to use their world knowledge in responding to the statements. In other words, they were told that some of the test statements should be considered true because the information in that statement was actually presented in the paragraph; other test statements should be considered true because the information described in them could be deduced from their existing knowledge about real things in the world (see also Potts, 1977).

The three sentences of a paragraph were presented simultaneously on the screen. However, as in Potts and Peterson (1985), they were not always presented in the standard order. Rather, the order of sentences was randomized for each paragraph and presented in the same randomized order to all participants. Participants controlled the display time for the three-sentence paragraph. When they had learned the paragraph adequately, they pressed the space bar to remove the paragraph from the screen, and the first test statement appeared. After participants responded to a test statement by pressing one of two keys labeled “yes” or “no,” the next test statement appeared. The 18 test statements were presented in a different random order for each participant. The dependent measure was the number of correct items for each type of statement.

Reading Comprehension Test

Participants were administered a standardized test of reading comprehension ability called the Nelson-Denny (Form E; Brown, Bennett, & Hanna, 1981). The Nelson-Denny consists of eight prose passages and 36 multiple-choice questions. Each multiple-choice question had five choices. Participants were given 20 minutes to read the passages and answer the questions.

Vocabulary Knowledge Test

Participants were administered the Mill Hill test of vocabulary knowledge. There were

20 multiple-choice items (e.g., fecund means [a] esculent, [b] profound, [c] sublime, [d] optative, [e] prolific, [f] salic). Participants completed all 20 items.

Verbal Analogies Test

Participants were administered a test of verbal analogies taken from the Inter-University Psychometric Entrance Exam (National Institute for Testing and Evaluation, 1991). There were 30 multiple-choice items (e.g., candle: lightbulb: [a] sun : moon, [b] arrow: bombshell, [c] floor: building, [d] battery: electricity). Participants had 10 minutes to answer as many items as possible.

Deductive Reasoning Test

Participants were administered the Poison Food deductive reasoning task (see Arenberg, 1968; Hayslip & Sterns, 1979). There were 10 problems. Each problem included 10 foods (e.g., beef, lamb, corn, peas, rice) and a number of clues (e.g., In the first meal beef, lamb, and corn were eaten and everyone died. In the second meal beef, peas, and rice were eaten and everyone died.). Participants were required to use the clues to deduce which food contained the poison (see Arenberg, 1968). Participants completed all 10 problems.

General Analytic Reasoning Test

Participants were administered 21 multiple-choice items based on the Graduate Record Examination analytic section (Educational Testing Service, 1992). Participants read a problem (e.g., Seven dignitaries, F, G, H, I, N, O and P are to be seated together at a diplomatic ceremony...) and a number of clues about that problem (e.g., F must sit in the chair immediately behind O's chair, N must sit in chair 3...) and answered three multiple-choice questions about that problem. They had 15 minutes to answer as many items as possible.

Results and Discussion

Table 1 presents the means and standard deviations on each component of the Potts and Peterson task, on the test of reading comprehension, and on the other verbal abilities tests. Table 2 shows the correlations among the four components of the Potts and Peterson task, and Table 3 shows how the components correlated with the five abilities tests. To allow for a direct comparison with the original Potts and Peterson (1985) task, the tables include the data based on the original four paragraphs as well as the data based on all eight paragraphs. The results of this experiment replicated Potts and Peterson's (1985) pattern of correlations among the four components of the task. However, no component was a particularly good predictor of reading comprehension ability. The poor predictive power of the Potts and Peterson task could not be attributed to the fact that the original version had only four paragraphs and 72 test statements; the pattern of results was almost identical when the number of paragraphs were doubled to eight and the number of test statements increased to 144.

Table 1
Means and Standard Deviations for Tasks Used in Experiment 1

Test	<u>M</u>	<u>SD</u>	<u>Range</u>
<u>Potts and Peterson's Task^a</u>			
<u>Original Four Paragraphs</u>			
Text Memory	84.4	17.6	33.0 - 100
Text Inferencing	81.1	21.4	37.5 - 100
Knowledge Integration	76.8	21.1	40.6 - 100
Knowledge Access	84.4	22.1	40.6 - 100
<u>All Eight Paragraphs</u>			
Text Memory	84.5	16.8	39.6 - 100
Text Inferencing	80.4	22.3	25.0 - 100
Knowledge Integration	79.3	19.4	37.5 - 100
Knowledge Access	87.2	18.7	12.5 - 100
<u>Tests of Reading and Verbal Abilities</u>			
Reading Comprehension (max = 36)	25.19	5.26	15 - 36 ^b
Vocabulary Knowledge (max = 20)	13.60	2.21	8 - 18
Verbal Analogies (max = 30)	12.26	4.62	1 - 23
Deductive Reasoning (max = 10)	7.18	3.41	0 - 10
General Analytic Reasoning (max = 21)	6.33	3.01	1 - 12

^aMeans, standard deviations, and ranges for the components in the Potts and Peterson (1985) task are reported as percentages.

^bAccording to the Nelson-Denny norms, this range represents the 7th to 100th percentiles for university students (Brown et al., 1981).

As Table 2 shows, the pattern of correlations among the four components of the Potts and Peterson (1985) task was replicated. For the original set of four paragraphs, the two text-based components, text memory and text inferencing, were highly correlated with one another (.80) and neither was correlated with knowledge access (.12 for text memory and .08 for text inferencing). In addition, knowledge integration, the component designed to measure the integration of prior knowledge with new text information, was significantly correlated with the two text-based components (.36 with text memory and .33 with text inferencing) as well as the knowledge access component (.67). And as Table 2 shows, a similar pattern emerged when the data for all eight passages were included in the correlational analyses. Thus, this replication experiment confirmed Potts and Peterson's (1985) finding that the ability to remember new text information and the tendency to access and use world knowledge are separate skills.

Table 2

Correlations Among the Four Components in the Potts and Peterson's Task Used in Experiment 1

	Text Inferencing	Knowledge Integration	Knowledge Access
Text Memory	.80*** (.86)***	.36** (.43)***	.12 (.15)
Text Inferencing	- - -	.33* (.50)***	.08 (.11)
Knowledge Integration		- - -	.67*** (.66)***

Note. Correlations based on Potts and Peterson's (1985) original four paragraphs are presented outside the parentheses; correlations based on all eight paragraphs are presented inside the parentheses.

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

However, as Table 3 shows, the four components of the Potts and Peterson (1985) task were at best only moderately correlated with reading comprehension ability, with correlations ranging from .16 to .38 for the original four paragraphs, and .23 to .32 for the complete set. A multiple regression analysis on the data for the original four paragraphs showed that the components accounted for 27% of the variance in reading comprehension (multiple $R = .52$), $F(3, 56) = 6.60$, $MSe = .77$, $p < .001$; a similar regression analysis on the data for the complete set of eight paragraphs showed that the components accounted for only 17% of the variance (multiple $R = .42$), $F(2, 56) = 5.70$, $MSe = 0.86$, $p < .006$.

Table 3
Correlations Between the Potts and Peterson Components and Other Verbal Abilities Tests
in Experiment 1 (n = 57)

	Reading Comprehension	Vocabulary	Verbal Analogies	Deductive Reasoning	Analytic Reasoning
Text Memory	.28* (.30*)	.12 (.15)	.15 (.17)	.39** (.44)***	.31* (.33)*
Text Inferencing	.38** (.32)*	.22 (.21)	.27* (.27)*	.36** (.35)**	.38** (.33)*
Knowledge Integration	.16 (.23)	.23 (.28)*	.15 (.19)	.44*** (.43)***	.37** (.46)***
Knowledge Access	.31* (.30)*	.35** (.35)**	.25 (.27)*	.38** (.38)**	.32* (.36)***
Reading Comprehension	---	.58***	.54***	.43***	.55***

Note. Correlations based on Potts and Peterson's (1985) original four paragraphs are presented outside the parentheses; correlations based on all eight paragraphs are presented inside the parentheses.

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

The four components correlated somewhat better with the deductive reasoning and general analytic reasoning tasks, with correlations ranging from .31 to .44 for the original four paragraphs, and from .33 to .46 for the complete set. Multiple regression analyses on the combined data for the two reasoning tasks showed that the components of the four-paragraph version accounted for 36% of the variance in reasoning, $F(3, 56) = 10.03$, $MSe = 1.88$, $p < .001$, and the components of the eight-paragraph version accounted for 38% of the variance in reasoning, $F(3, 56) = 10.81$, $MSe = 1.82$, $p < .001$. It is not surprising that the Potts and Peterson task was quite a good predictor of performance on the reasoning tasks given that it was specifically designed to engage the problem-solving skills for constructing a five-term linear ordering and reasoning about the relations among the five terms (see Potts, 1977). But why was the task such a disappointing predictor of reading comprehension performance?⁴ After all, on the basis of previous research findings, it seems highly plausible to assume that memory, inferencing, knowledge access, and knowledge integration are important components of reading comprehension.

Upon reflection, it appears that the main problem with the Potts and Peterson (1985) task was that it is just not complex enough to capture the kinds of memory, inferencing, knowledge access, and knowledge integration processes that are part and parcel of full-blown reading comprehension. Indeed, almost 60% of the participants reported performing the task by memorizing a simple mnemonic for the five-term linear ordering (e.g., JTPBC for JAL > TOC > PONY > BEAVER > CAZ) and rehearsing the mnemonic throughout the testing phase. Although

⁴See also Singer et al. (1992), who found a low .29 correlation between Potts and Peterson's knowledge integration component and performance on the Nelson-Denny test of reading comprehension ability.

developing and rehearsing a mnemonic may be a useful strategy for performing well on the Potts and Peterson (1985) task, it is unlikely that such a strategy plays a major role in reading comprehension. And so even though part of the original appeal of the Potts and Peterson task was its simplicity, I recognized that I might need to increase the complexity of the task to increase its predictive power. This was accomplished by increasing the complexity of both the three-sentence paragraphs and the test statements. In Chapter 4, the modified version of the Potts and Peterson four-component task is described, and the extent to which the four components are able to predict individual differences in reading comprehension ability is examined.

Chapter 4: Experiment 2

The first strategy for increasing the complexity of the Potts and Peterson (1985) task was to increase the complexity of the three-sentence paragraphs. The Potts and Peterson paragraphs could be represented as simple five-item linear orderings because they included only one relation among the five terms (e.g., larger than). I increased the complexity of representing and reasoning about the paragraphs by including two to four relations/features. Consider the following paragraph:

A MIRT resembles an OSTRICH but is larger and has a longer neck.

A COFT resembles a ROBIN but is smaller and has a longer neck.

A FILP resembles a COFT but is smaller, has a longer neck, and nests on land.

As in the Potts and Peterson task, these new paragraphs contained three nonsense terms (e.g., MIRT, COFT, and FILP) and two real terms (e.g., OSTRICH, ROBIN). However, unlike the Potts and Peterson task which only included one feature, the new paragraphs included two to four features (e.g., larger than, longer neck, and nests on land). The additional features create a complex structure. In this example, the size feature produces the linear ordering MIRT > OSTRICH > ROBIN > COFT > FILP if individuals access their prior knowledge that an ostrich is larger than a robin. The neck feature produces two orderings, MIRT > OSTRICH > ROBIN, and FILP > COFT > ROBIN if individuals access their prior knowledge that an ostrich has a longer neck than a robin. The third feature, nests on land, adds to the complexity of the representation and provides additional ways to test learning.

The second strategy for increasing the complexity of the Potts and Peterson (1985) task

was to increase the number and types of test statements. By having more than one semantic feature (e.g., larger than, neck length, and nests on land), the number of text memory and text inferencing statements per paragraph could be increased. For example, there were text memory statements testing the larger than relation (e.g., A MIRT is larger than an OSTRICH), the neck length relation (e.g., A MIRT has a longer neck than an OSTRICH), and the nests on land relation (e.g., A FILP nests on land). Similarly, there were text inferencing statements testing all three relations (e.g., A FILP is smaller than a ROBIN; A FILP has a longer neck than a ROBIN; A COFT doesn't nest on land). A second type of knowledge access statement was also included. The low-knowledge access statements (e.g., An OSTRICH is larger than a ROBIN) were like those used in the Potts and Peterson (1985) task in that they used terms (e.g., OSTRICH, ROBIN) and a semantic feature (e.g., larger than) explicitly mentioned in the paragraph. On the other hand, the new high-knowledge access statements (e.g., A ROBIN lives in Canada, whereas a PENGUIN typically doesn't) required more extensive use of prior knowledge because they used a term explicitly mentioned in the paragraph (e.g., ROBIN), and a term (e.g., PENGUIN) and semantic feature (e.g., lives in Canada) not explicitly mentioned in the paragraph. Finally, two new types of knowledge integration statements were included. The low-knowledge integration statements (e.g., A MIRT has a longer neck than a ROBIN) were like the knowledge integration statements used by Potts and Peterson (1985) in that they tested knowledge integration with two terms (e.g., MIRT and ROBIN) and a semantic feature (e.g., longer neck) all explicitly mentioned in the paragraph. The medium-knowledge integration statements (e.g., A MIRT is larger than a BLUEJAY) tested integration of information explicitly mentioned in the paragraph (e.g., MIRT, larger than) with a term not explicitly mentioned in the paragraph (e.g.,

BLUEJAY). Finally, the high-knowledge integration statements (e.g., Like PENGUINS, MIRTS can't fly) tested integration of a nonsense term mentioned in the paragraph (e.g., MIRT) with a semantic feature (e.g., can't fly) and a term (e.g., PENGUIN) not explicitly mentioned in the paragraph. So, in summary, the new task included more text memory statements, more text inferencing statements, two types of knowledge access statements (low, high), and three types of knowledge integration statements (low, medium, high).

By increasing the complexity of the paragraphs with additional semantic features and by increasing the number and complexity of test statements, the new task should tap more complex memory, inferencing, knowledge access, and knowledge integration processes than does the Potts and Peterson (1985) task, and thus should be more likely to capture the kinds of processes involved in full-blown reading comprehension. This possibility was investigated in Experiment 2 by administering the new component processes task in conjunction with the same test of reading comprehension ability administered in Experiment 1 (Chapter 3).

Although the major focus of Experiment 2 was to determine the extent to which the components of the new component processes task predicted performance on a standardized test of reading comprehension ability, the vocabulary and verbal reasoning tests used in Experiment 1 were included. In addition, a nonverbal spatial rotation test was added to the test battery. The goal was to provide both convergent and discriminant validation for the component processes task by showing that the task predicted performance on the verbal abilities tests but did not predict performance on the spatial abilities test.

Method

Participants

The participants were 69 University of Toronto students who had not participated in Experiment 1. All participants were fluent speakers of English and were tested individually in two sessions. Participants were administered the new component processes task followed by the same five tests of verbal abilities used in Experiment 1 (reading comprehension, vocabulary knowledge, verbal analogies, deductive reasoning, and general analytic reasoning) as well as a test of spatial rotation ability. Below only the new tasks are described.

Component Processes Task

Participants studied three-sentence paragraphs such as the “bird” one, and then answered a set of true-false statements about the paragraph.

Materials. The materials consisted of a set of seven short paragraphs. The first paragraph was used as a practice paragraph. As in the Potts and Peterson (1985) task used in Experiment 1, the new paragraphs each consisted of three sentences. However, the sentences were not presented simultaneously on the computer screen, nor were they scrambled. Rather, each sentence appeared one at a time and in the same standard order. Each paragraph included three nonsense terms, two real terms, and either two, three, or four semantic features. For example, in the following paragraph,

A NORT resembles a JET but is faster and weighs more.

A BERL resembles a CAR but is slower and weighs more.

A SAMP resembles a BERL but is slower and weighs more.

NORT, BERL, and SAMP are the nonsense terms, JET and CAR are the real terms, and speed

and weight are the two semantic features. The number of features per paragraph increased every two paragraphs. Thus, the first two experimental paragraphs had two features, the next two had three features, and the final two had four features. As was the case for the Potts and Peterson (1985) paragraphs, participants had to use their existing knowledge to construct the linear orderings. For example, in the “vehicle” paragraph above, the fact that a jet is faster than a car and the fact that a jet weighs more than a car are not explicitly mentioned, and so participants need to use their existing knowledge of these facts to construct the speed and weight linear orderings. To ensure that all the prior knowledge facts used in the stimuli were part of a university student’s existing knowledge, an independent group of 67 University of Toronto undergraduates were tested on these facts. All facts were known by 98 % of the undergraduates, which suggests that the prior knowledge facts used in the paragraphs are indeed part of a university student’s existing knowledge. See Appendix A for the complete set of experimental paragraphs.

After studying a paragraph, participants responded to true--false statements about it. Half of the statements were true and the other half were false. False statements were created by reversing the terms in the true statements. There were four main types of test statements: text memory, text inferencing, knowledge access, and knowledge integration. In all, participants responded to 40 test statements for the practice passage and 276 test statements for the six experimental passages. See Appendix A for the complete set of test statements.

The text memory statements tested memory for information explicitly presented in the paragraph; no prior knowledge was required (e.g., A NORT is faster than a JET). The total number of text memory statements for a paragraph depended on the number of semantic features

in that paragraph. Paragraphs with two features had six true and six false text memory statements; paragraphs with three features had seven true and seven false text memory statements; paragraphs with four features had eight true and eight false text memory statements.

The text inferencing statements tested inferences about information presented explicitly in the paragraph; no prior knowledge was required (e.g., A SAMP is slower than a CAR which can be inferred from the text facts A BERL is slower than a CAR and A SAMP is slower than a BERL). Paragraphs with two semantic features had two true and two false text inferencing statements; paragraphs with three semantic features had three true and three false text inferencing statements, and paragraphs with four semantic features had four true and four false text inferencing statements.

The knowledge access statements tested access to prior knowledge; no information from the paragraph was required. There were two types of knowledge access statements: low and high. Low-knowledge access statements (e.g., A JET is faster than a CAR) tested access to a fact not presented in the paragraph but included two real terms (JET and CAR) and a feature (faster than) that appeared in the paragraph. Paragraphs with two semantic features had two true and two false low-knowledge access statements; paragraphs with three features had three true and three false low-knowledge access statements, and paragraphs with four semantic features had four true and four false low-knowledge access statements. High-knowledge access statements (e.g., A JET has a pilot, whereas a MOTORCYCLE doesn't) tested access to a fact not presented in the paragraph and included a real term (JET) that was presented in the paragraph, as well as a real term (MOTORCYCLE) and semantic feature (has a pilot) not presented in the paragraph. By including a term and a feature not explicitly mentioned in the paragraph, the high-knowledge

access statements required more sophisticated access to and reasoning about the relations among sources of prior knowledge. For each paragraph, there were two true and two false high-knowledge access statements.

Finally, the knowledge integration statements tested integration of prior knowledge with text information. There were three types of knowledge integration statements that increased in level of difficulty: low, medium, and high. Low-knowledge integration statements (e.g., A NORT weighs more than a CAR) required participants to access their prior knowledge that a jet weighs more than a car and integrate this fact with the text information that A NORT weighs more than a JET. These were like the Potts and Peterson (1985) knowledge integration statement in that each statement included a nonsense term (NORT), a real term (CAR), and a feature (weighs more) that all appeared in the paragraph. However, by having two linear ordering terms (e.g., faster than, weighs more), I was able to increase the number of low-integration statements by testing the integration of prior knowledge for both linear orderings (see Appendix A for examples). The medium-knowledge integration statements (e.g., A NORT is faster than a MOTORCYCLE) also required participants to access prior knowledge and integrate it with text information; however, the nonsense term (NORT) and the feature (faster than) appeared in the paragraph, whereas the real term (MOTORCYCLE) did not. The high-knowledge integration statements (e.g., Like ROCKETS, NORTS travel in the air) also required participants to access prior knowledge and integrate it with text information, and each included a nonsense term (NORT) that appeared in the paragraph, but neither the real term (ROCKETS) nor the semantic feature (travels in the air) appeared in the paragraph. Each paragraph had two true and two false low-knowledge integration statements, three true and three false medium-integration statements,

and three true and three false high-integration statements.⁵

Procedure. As in Experiment 1, participants were explicitly instructed to use their world knowledge in performing the task (See Appendix B for the full instructions). The instructions and procedure were identical to those for the Potts and Peterson (1985) task used in Experiment 1, except that the three sentences of a paragraph were presented one at a time in the standard order.⁶ Participants controlled the display time for a sentence by pressing the '+' key when they were ready for the next sentence. After a participant had learned all three sentences of a paragraph, the test statements for that paragraph were presented one at a time in a random order. Participants had a maximum of 12 seconds to read and respond to a test statement. If the participant failed to respond within the 12-second deadline, the test statement was removed and replaced by the next test statement. All response failures were classified as errors.⁷ Accuracy (i.e., number correct) was the primary dependent measure; however, speed of responding (i.e., average reaction time for correct responses) was also analyzed. A pause screen appeared at the end of the test statements for each paragraph to give participants a break before proceeding to the next paragraph.

⁵Each true high-knowledge integration statement (e.g., Like ROCKETES. NORTS travel in the air) had a corresponding false high-knowledge integration statement (e.g., Like MOTORCYCLES. NORTS travel across the land) for which both the new term (e.g., ROCKETES) and the feature (e.g., travels in air) were changed.

⁶Potts, Keller, and Rooley (1981) found no differences between a scrambled presentation order of the sentences in the paragraph and an unscrambled presentation order. Because I was interested in making my task more complex, I presented the sentences in a standard order but removed a sentence from the screen after a participant pressed the '+' key for the next paragraph sentence.

⁷Response failures as a result of time-outs accounted for approximately 1% of the trials in this and subsequent experiments.

Spatial Rotation Ability Test

This test was an adaptation of Shepard and Metzler's (1971) spatial rotation test. There were 20 items. Each item consisted of a drawing that depicted a complex three-dimensional object made up of small blocks, followed by four more drawings of multiblock objects, two of which depicted the same target object but drawn at a different angle of orientation. The task was to identify the two drawings that depicted the target object. Participants were given five minutes to complete as many items as possible.

Results and Discussion

Table 4 presents the means and standard deviations on each component of the new component processes task and on the six abilities tests. Table 5 shows the correlations among the components of the task for the accuracy data, and Table 6 shows how the components correlate with the six abilities tests. As predicted, all the components of the new component processes task were more highly correlated with global reading comprehension ability than were the components of the Potts and Peterson task used in Experiment 1.

Table 4

Means and Standard Deviations for Tasks Used in Experiment 2 (n = 69)

Test	<u>Accuracy</u>		<u>Speed</u>	
	<u>M/(SD)</u>	<u>Range</u>	<u>M/(SD)</u>	<u>Range</u>
<u>Component Processes Task^a</u>				
Text Memory	74.69 (11.86)	42.86 - 98.81	4008(803)	2380 - 5582
Text Inferencing	67.03 (13.62)	33.33 - 94.44	4003(947)	2314 - 6323
Low-Knowledge Integration	84.78 (12.08)	54.17 - 100	3557(815)	1855 - 5350
Medium-Knowledge Integration	80.64 (12.88)	50.00 - 100	3761(838)	1848 - 5751
High-Knowledge Integration	71.91 (14.55)	41.67 - 97.22	4382(926)	2458 - 6370
Low-Knowledge Access	91.87 (6.13)	69.44 - 100	3296(620)	2241 - 4558
High-Knowledge Access	86.90 (9.35)	54.17 - 100	4424(857)	2733 - 6389
<u>Tests of Reading and Verbal Abilities</u>				
Reading Comprehension (max = 36)	24.32 (5.63)	11 - 33 ^b		
Vocabulary Knowledge (max = 20)	13.77 (2.01)	8 - 17		
Verbal Analogies (max = 30)	12.51 (5.11)	3 - 26		
Deductive Reasoning (max = 10)	7.29 (3.27)	0 - 10		
General Analytic Reasoning (max = 21)	5.87 (2.85)	1 - 13		
<u>Test of Nonverbal Abilities</u>				
Spatial Rotation (max = 40)	15.02 (6.59)	5 - 39		

^aMeans, standard deviations, and ranges for the components are reported as percentages.

^bAccording to the Nelson-Denny norms, this range represents the 2nd to 93rd percentiles (Brown et al., 1981) for university students.

The components of the new task showed a similar pattern of intercorrelations as seen in the Potts and Peterson (1985) task used in Experiment 1. As Table 5 shows, the two text-based components, text memory and text inferencing, were highly correlated with one another (.83), and both were at best weakly correlated with the knowledge access components (correlations ranging from .18 to .30). On the other hand, all three knowledge-integration components were significantly correlated with the two text-based components (correlations ranging from .54 to .70) and they were also correlated with the knowledge access components, although the correlations were higher for the low-knowledge access component (correlations ranging from .34 to .49) than for the high-knowledge access component (correlations ranging from .22 to .31). Taken together, these findings suggest that the two text-based components are tapping different skills than are the knowledge access components, but that the knowledge integration components are tapping both prior knowledge and text knowledge.⁸ Finally, as expected, the three knowledge integration components were strongly correlated with one another (correlations ranging from .53 to .72), although the correlation between the two measures of knowledge access was weaker than expected (.32; but see Experiment 3, Chapter 5).

⁸Note that in the Potts and Peterson (1985) task, knowledge integration was more highly correlated with knowledge access than with the text-based components, text memory and text inferencing. In the new task, knowledge integration was more highly correlated with the text-based components in this and subsequent experiments.

Table 5

Correlations Among the Components of the Component Processes Task in Experiment 2 (n = 69)

	Text Inferencing	Low- Knowledge Integration	Medium- Knowledge Integration	High- Knowledge Integration	Low- Knowledge Access	High- Knowledge Access
Text Memory	.83***	.60***	.69***	.70***	.30*	.18
Text Inferencing	- - -	.54***	.58***	.61***	.28*	.27*
Low- Knowledge Integration		- - -	.72***	.53***	.42**	.31*
Medium- Knowledge Integration			- - -	.70***	.49***	.27*
High- Knowledge Integration				- - -	.34**	.22
Low- Knowledge Access					- - -	.32**

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

Reliability estimates were calculated for each component using Cronbach's alpha formula as a measure of internal consistency. This involved calculating two scores for each participant on each component, one that was based on the items for one set of two-, three-, and four-feature paragraphs, and a second that was based on the items for the other set of two-, three-, and four-feature paragraphs, and then computing the correlations between these aggregated scores. The reliability coefficients for all the components were high, ranging from .86 to .88.

The correlations among the components based on speed of responding were also computed. The pattern differed from the pattern for the accuracy measures in the sense that the speed of responding correlations among the components were all high, ranging from .55 to .86. These across-the-board high correlations, together with the fact that none of the speed measures was significantly correlated with its corresponding accuracy measure (mean correlation was .16, mean $p = .26$) suggest that the speed measures were likely tapping some common factor to do with speed of reading and responding to a test statement and were not sensitive to the particular component processes that an individual test statement was designed to measure.⁹ Consequently, only one speed measure was included in the correlational and regression analyses to follow. For the purposes of this thesis, I call this component, speed.¹⁰ Cronbach's alpha reliability estimate for the speed component was .85.

Table 6 shows that all the components of the new component processes task were

⁹This lack of correlation also suggests that there are no speed-accuracy trade-offs. See Jackson & McClelland, 1979, pp. 179 for an example of a speed-accuracy trade-off.

¹⁰I chose the speed measure that correlated highest with reading comprehension ability. This was the speed of responding to high-knowledge access statements. This is also the speed measure reported and analyzed in Experiments 3 (Chapter 5) and 4 (Chapter 6).

significantly correlated with performance on the Nelson-Denny test of reading comprehension ability, with correlations ranging from .36 to .62.¹¹ These results contrast quite strikingly with the results for the simpler Potts and Peterson task used in Experiment 1; for that task, the highest correlation between a component and reading comprehension ability was .38. Notice also that the new task did a reasonably good job of predicting performance on some of the verbal reasoning tasks. However, with the exception of what has been interpreted to be a general speed factor, none of the components of the new task correlated with performance on the test of spatial rotation ability. This finding is consistent with previous research that argues for a dissociation between verbal and spatial abilities (e.g., Daneman & Tardif, 1987; Kyllonen, 1993; Shah & Miyake, 1996) and provides some discriminant validity for the new measure.

¹¹Analyses revealed that there was no evidence of nonlinear relationships between reading comprehension ability and any of the component processes that were used as predictors of reading comprehension ability. This was also the case in Experiments 3 (Chapter 5) and 4 (Chapter 6).

Table 6
Correlations Between Components of the Component Processes Task and Abilities Tests
in Experiment 2 (n = 69)

	Reading Comprehension	Vocabulary	Analogies	Deductive Reasoning	Analytic Reasoning	Spatial Rotation
Text Memory	.43***	.26*	.37**	.27*	.48***	-.09
Text Inferencing	.45***	.25*	.33**	.33**	.34**	-.13
Low-know Integration	.36**	.13	.22	.35**	.36**	-.05
Med-know Integration	.40***	.23	.45***	.35**	.50***	.07
High-know Integration	.62***	.29*	.46***	.23	.44***	-.05
Low-know Access	.45***	.24*	.33**	.35**	.27*	.14
High-know Access	.37**	.22	.23	.14	.09	.19
Speed	-.46***	-.24*	-.20	-.24*	-.15	-.34**
Reading Comprehension	---	.47***	.30*	.32**	.45***	.11

* p < .05, ** p < .01, *** p < .001

From the individual correlations in Table 6, it is evident that the high-knowledge integration component was the best single predictor of reading comprehension ability, accounting for 38.4% of the variance in performance on the Nelson-Denny test of reading comprehension ability. I would argue that the high-knowledge integration component is a better predictor of reading ability than are the simple text-based components or the simple knowledge access components because it is a complex component that draws not only on text-based processes and knowledge access processes, but also on the processes involved in integrating the text information with prior knowledge. I would argue that high-knowledge integration is also a better overall predictor of reading ability than are the two other knowledge integration components (low- and medium-knowledge integration), because it likely taps more sophisticated knowledge access and integration processes by making individuals draw on knowledge about a concept and a feature not explicitly mentioned in the studied paragraphs. Presumably, these richer knowledge access and integration processes would be part and parcel of the kinds of thematic and predictive knowledge-based inferences that skilled readers make to embellish their text representations (Hannon & Daneman, 1998; Long et al., 1994).

Of course, also of interest was the amount of variance in reading comprehension performance that could be accounted for by the task as a whole. Consequently, a series of multiple regression analyses were conducted to address this question. Following Dixon et al. (1988), my strategy was to test for the unique contribution of a component process by entering it into the regression equation last. If that component process failed to make a unique contribution,

it was dropped from the regression equation.¹² Regression analyses with the remaining components showed that the component processes task could account for 60% of the variance in performance on the Nelson-Denny test of reading comprehension (multiple $R = .774$). This 60% is much higher than the 27% accounted for by the Potts and Peterson (1985) task used in Experiment 1, $z = 2.49$, $p = .0064$. The result is particularly impressive when the reliability of the Nelson Denny test is taken into consideration; indeed, Form F of the Nelson-Denny accounts for only 59% of the variance in Form E, the form used in this study (Brown et al., 1981). Thus, it appears that the component processes task may be as good at predicting reading comprehension ability as is another test of reading comprehension ability. Moreover, there is the added theoretical bonus that the task provides some insight into the nature of the component processes tapped by an atheoretical global reading comprehension test such as the Nelson-Denny. Finally, there is the added psychometric bonus that the task provides a practical tool for measuring individual differences in these theoretically important component processes.

Table 7(a) shows the results of one of the regression analyses. The procedure adopted for this regression model was as follows: After establishing which components made unique contributions to reading comprehension when added into the regression equation last, these components were then classified as either elementary (low-knowledge access, high-knowledge access, text memory, text inferencing, speed) or complex (low-, medium-, and high-knowledge integration). Elementary components were entered into the equation before complex components

¹²Because there was such a high correlation between text memory and text inferencing, both of these components failed to make a unique contribution to the equation when entered last. To handle this problem, the text-based component that contributed the most variance was used when it was the second last component entered into the model.

on the assumption that more complex component processes (e.g., knowledge integration) are dependent on the presence of more elementary component processes (e.g., text memory, low-knowledge access). I am not wedded to the particular model of reading comprehension presented in Table 7(a), but offer it merely as an illustration of the pattern and proportion of variance that can be accounted for by the components of the new task. As Table 7(a) shows, the elementary text-based, knowledge access, and speed components accounted for 49% of the variance in reading comprehension performance. However, the complex knowledge integration component which draws on text based processes, knowledge access processes, and the integration of the two, still accounted for a significant proportion of unique variance (11%) over and above the elementary components of which it is composed.

Regardless of the order in which the predictors were entered into a regression model, the total variance accounted for by the regression model remained 60%. Moreover, after a predictor was added into a model, that predictor's variance did not change when subsequent predictors were added. Regression analyses were also conducted in which the low- and high-knowledge access components were combined into a single knowledge access predictor rather than allowing them to enter into the analysis as two separate predictors, and the low-, medium-, and high-knowledge integration components were combined into a single knowledge integration predictor rather than allowing them to enter into the analysis as three separate predictors. The regression equations using aggregated knowledge access and knowledge integration scores were similar to those using separate predictors, but accounted for 1% to 2% less total variance.

Table 7
Regression Analyses on Reading Comprehension Test Scores Using the Components
of the Component Processes Task as Predictors (n = 69)

Variable	R	R ²	ΔR^2	F
(a) Reading comprehension				
1. Speed	.46	.21	.21	17.81
2. Text inferencing	.64	.41	.20	22.82
3. Low-knowledge access	.68	.47	.06	18.84
4. High-knowledge access	.70	.49	.02	15.01
5. High-knowledge integration	.77	.60	.11	18.03
(b) Vocabulary as first predictor of reading comprehension				
1. Vocabulary	.47	.22	.22	18.85
2. Speed	.59	.35	.13	12.68
3. Text inferencing	.69	.48	.13	16.68
4. Low-knowledge access	.72	.52	.04	5.48
5. High-knowledge access	.73	.53	.01	1.73
6. High-knowledge integration	.79	.63	.10	16.40
(c) Vocabulary as predictor of reading comprehension after variance accounted for by the component processes has been partialled out				
1. Speed	.46	.21	.21	17.81
2. Text inferencing	.64	.41	.20	22.82
3. Low-knowledge access	.68	.47	.06	18.03
4. High-knowledge access	.70	.49	.02	15.01
5. High-knowledge integration	.77	.60	.11	18.81
6. Vocabulary	.79	.63	.03	5.34

Note. ΔR^2 = change in R^2 .

Vocabulary knowledge has been considered a good, if not the best, predictor of reading comprehension ability (see Daneman, 1988, 1991; Dixon et al., 1988; Just & Carpenter, 1987; Sternberg & Powell, 1983; Thorndike, 1973). Thus, I was interested in evaluating the extent to which the new component processes task and the commonly used Mill Hill test of vocabulary knowledge were making independent or overlapping contributions toward predicting reading comprehension performance. To do this, two regression analyses were conducted, one in which vocabulary knowledge was entered as the first predictor and a second in which it was entered after speed, text inferencing, low-knowledge access, high-knowledge access, and high-knowledge integration were partialled out. As Table 7(b) shows, when entered first, vocabulary accounted for 22% of the variance in reading comprehension performance, a finding that is consistent with previous research (e.g., Dixon et al., 1988); however, note also that the component processes still accounted for a substantial 41% of the variance in reading comprehension after the effects of vocabulary knowledge were partialled out statistically. On the other hand, as Table 7(c) shows, when vocabulary was entered into the regression equation after the effects of the component processes had been statistically removed, vocabulary accounted for only 3% of unique variance. Thus, it appears that the new component processes task is capturing most of the variance in reading comprehension that is tapped by a typical test of vocabulary knowledge.

In summary, the results of Experiment 2 are promising with respect to the predictive validity of the new component processes task because they show that it accounts for a considerable proportion of the variance in performance on a standardized test of global reading comprehension ability. The experiments reported in Chapters 5 and 6 provide evidence for the

reliability of these findings by replicating the basic pattern of intercorrelations among the components as well as the correlations between the components and the Nelson-Denny test of reading comprehension ability.

Chapter Five: Experiment 3

Experiment 3 had two goals. The first goal was to replicate the Experiment 2 findings with respect to the basic pattern of intercorrelations among the components of the new task and the ability of the components to predict performance on a global test of reading comprehension ability. The second goal was to investigate how the components of the new task predicted performance on a range of specific tests of reading comprehension that were designed to load on one or more of the specific components, namely, text memory, text inferencing, knowledge access, and knowledge integration.

Examining how the new task correlated with specific tests of reading comprehension served two functions. First, it was important to show that the predictive power of the new task is not limited to one test of reading ability, in this case, the Nelson-Denny. The Nelson-Denny was selected as a criterion task because it is a widely-used and standardized test of reading comprehension ability. However, it would be a mistake to equate a complex construct such as reading comprehension ability with performance on any single measure of reading comprehension. By including a variety of reading comprehension tasks that differed in the nature and length of the texts-to-be-read, as well as in the way in which comprehension was assessed, I could show that the predictive power of the task is not limited to a multiple-choice reading comprehension test such as the Nelson-Denny. But even more important, I wanted to provide additional construct validity for each component of the new task. In other words, I wanted to obtain more evidence that the text memory component is indeed measuring text memory, that the text inferencing component is indeed measuring text inferencing, and so on. So

far, I have relied on the face validity of the task and on the pattern of intercorrelations among the components to make this argument. In other words, I have relied on the following findings: text memory statements, which appear to depend only on text-based processes, are correlated best with the text inferencing statements, the other type of test statement that appears to depend only on text-based processing; both text memory and text inferencing statements show intermediate correlations with knowledge integration statements, the test statements that appear to depend on text-based processes as well as knowledge access; both text memory and text inferencing statements are correlated most weakly with knowledge access statements, the test statements that appear to involve only knowledge access. The rationale for including specific tests of reading comprehension that load on one or more of the specific components of reading was that I could provide external validation for the components of the new task if I found that the text-memory component was the best predictor of performance on a specific reading task that taxes memory for text details, that the text inferencing component was the best predictor of performance on an inference-loaded reading task, and so on.

Of course, I recognized from the outset that this would be a very tricky enterprise because as soon as the reading task is made sufficiently complex, all the important components of reading tapped by the task--memory, inferencing, knowledge access, and knowledge integration--will likely be engaged. So to avoid this potential outcome, my strategy was to choose some very simple reading tasks that involved short texts and specific modes of testing that were each likely to draw more heavily on one particular component. However, I also included two more complex reading tasks that were likely to engage numerous component processes. Although I made predictions about which component(s) would best predict performance on these more complex

reading tasks, my main interests here were in the relative contributions of the individual components and the extent to which these patterns differed from the pattern found for the standardized test of global reading comprehension.

Method

Participants

There were 81 participants: 22 had taken part in Experiment 2 and 59 were new participants who had not taken part in the previous experiments. All participants were fluent speakers of English and were tested individually. The 59 new participants were administered the identical component processes task and Nelson-Denny test of reading comprehension that the other 22 had been administered in Experiment 2. Then all 81 participants were administered five specific tests of reading comprehension. The three simple reading tasks were (a) a short-stories task (Oakhill, 1982) that loaded on text memory and text inferencing; (b) a verification task (Smith, Shoben, & Rips, 1974) that loaded on knowledge access; and (c) a bridging-inference task (Keenan & Kintsch, 1974) that loaded on knowledge integration. The two more complex specific reading tasks tapped (a) the ability to infer the meanings of new words from context (Daneman & Green, 1986), and (b) memory for text details. Below, the five specific tests are described.

Short-Stories Task

This was a variation of a task used by Oakhill (1982) in which participants read eight three-sentence stories (e.g., The plane flew over the house. The house was in Crawley. The plane landed in a field.) and then after a five-minute distractor task, responded to true/false statements of two types: one type tested memory for explicit text details (e.g., The plane flew

over the house.), and another type tested simple inferences (e.g., The plane flew over Crawley.).

I predicted that the text memory component of the component processes task should be the best predictor of performance on Oakhill's (1982) memory statements because these statements tested information mentioned explicitly in the stories; no inferencing or access to prior knowledge was required. In contrast, I predicted that the text inferencing component would be the best predictor of performance on Oakhill's (1982) inference statements because these statements required the reader to go beyond the explicit text and make an inference based largely on information mentioned explicitly in the passage.

The eight stories and accompanying test statements were taken from Oakhill (1982). Participants were instructed to read each story carefully because their memory for the stories would be tested later. The stories were presented one sentence at a time in the middle of a computer screen. Once a participant pressed the '+' key for the next sentence, the previous sentence was removed from the screen. If the participant did not press the '+' key for the next sentence within 10 seconds, the current sentence was removed and the next sentence appeared. Participants read all eight short stories and then performed a five-minute distractor task, namely a word-stem completion task.

During the test phase, participants answered true-false statements and were instructed to press the 'yes' key if the statement was either identical to one of the sentences in a story or was implied by a story, and to press the 'no' key if it was not identical or implied. There were four test statements for each story. Two were true memory statements (e.g., The plane flew over the house.), one was a true inference statement (e.g., The plane flew over Crawley.), and one was a false inference statement (e.g., The house was in a field.). Each statement was presented

individually in the center of the computer screen. All statements pertaining to the first story were presented first in a randomized order, then all statements pertaining to the second story were presented in a randomized order, and so on. All responses were recorded by the computer.

Verification Task

This was an adaptation of a task used by Smith et al. (1974; see also Shoben, 1992) in which participants read and verified simple statements such as An apple is a fruit., A cantaloupe is a fruit., An apple is a vegetable. I predicted that the real-knowledge component of the component processes task would be the best predictor of performance on this task because the verification task required access to prior knowledge about semantic category membership and made little if any demands on the text-based and integration processes of reading.

All critical statements involved the fruit category (e.g., An apple is a fruit.). Fruit exemplars were either highly typical (e.g., apple), moderately typical (e.g., cantaloupe), or atypical (e.g., cranberry). On the basis of Rosch's (1975) typicality ratings, highly typical exemplars had ratings from 1.00 to 1.75; moderately typical exemplars had ratings from 1.76 to 2.99; and atypical concepts had ratings greater than 2.99. For every true statement there was a corresponding false statement in which the category label fruit was substituted with vegetable (e.g., An apple is a vegetable.). In total, there were 21 true statements and 21 false statements--14 for each of the three typicality levels.

The filler statements were about the vegetable category (e.g., A pea is a vegetable.). All vegetable exemplars were classified as highly typical (e.g., pea), moderately typical (e.g., turnip), or atypical (e.g., mushroom). The ranges for highly typical, moderately typical, and atypical ratings were identical to those used for the critical statements. For every true statement there was

a corresponding false statement in which the category label vegetable was substituted with fruit (e.g., A pea is a fruit.). In total, there were 21 true and 21 false filler statements—14 for each of the three typicality levels.

Each statement was presented individually in a random order in the center of a computer screen. Participants were asked to respond as quickly as possible but not so quickly as to make errors. They pressed the ‘yes’ key if they believed the statement was true, and the ‘no’ key if they believed the statement was false. Participants responses and reaction times were recorded by the computer.

Bridging Inference Task

This was an adaptation of Keenan and Kintsch’s (1974) task in which participants read short passages that did or did not require a bridging inference. For example, the implicit passage, A burning cigarette was carelessly discarded. The fire destroyed many acres of virgin forest., requires a bridging inference that ‘the cigarette started the fire’ to make it coherent. In other words, it requires readers to access their prior knowledge that cigarettes can cause fires and to integrate this knowledge with text information to make sense of the text. In contrast, the explicit passage, A carelessly discarded cigarette started a fire. The fire destroyed many acres of virgin forest., does not require readers to access their prior knowledge and make the bridging inference because the text explicitly states that the cigarette was the cause of the fire. Fifteen minutes after reading a series of implicit and explicit passages, participants responded to whether test statements such as A cigarette started the fire were true based on information either implied or stated in one of the passages they had read. Keenan and Kintsch (1974) found no difference in reaction time to verify implicit versus explicit statements. On the basis of this finding, they

argued that readers make bridging inferences when reading implicit passages and that the representation they store for implicit passages is equivalent to that stored for explicit passages. Of course, Keenan and Kintsch analyzed group data only. It is quite possible that readers might differ in their ability or tendency to make bridging inferences. Given this possibility, I predicted that the knowledge integration components would be the best predictors of these inferences because these inferences involved accessing prior knowledge and integrating this knowledge with text information (see also Singer et al., 1992). Specifically, I predicted that readers who were skilled on the knowledge integration components would be more accurate at verifying implicit statements than would readers who were less skilled on the knowledge integration components. Moreover, skilled integrators should also show smaller reaction time differences for implicit minus explicit statements, suggesting that they had indeed made the knowledge-based inferences.

The passages and test statements were taken from or modelled after those used by Keenan and Kintsch (1974). There were three types of passages: implicit, explicit, and filler. The 26 implicit passages required a bridging inference (e.g., 'The cigarette started the fire') to make them coherent, whereas the 26 explicit counterparts provided the information explicitly. The 26 filler passages were similar to the explicit passages in that they were two sentences in length and did not require a bridging inference to make them coherent. The average lengths of the implicit, explicit, and filler texts were 15.70, 15.77, and 17 words, respectively. There were 26 true test statements that corresponded to the implicit and explicit passages (e.g., A cigarette started the fire) and 26 false statements that corresponded to the filler passages. Because each participant saw only the implicit or explicit version of a passage, 13 of the true test statements corresponded

to implicit passages and 13 corresponded to explicit passages.

Participants were told that they would be reading a series of passages, each two sentences in length. They were told that the two sentences should be integrated as if they were a story (see Keenan & Kintsch, 1974). Each passage was presented one at a time on a computer screen. Participants controlled the presentation duration by pressing the '+' key for the next passage. After reading all 52 passages, there was a 15-minute distractor task, and then participants responded to the 52 test statements. Test statements were presented individually in a random order in the center of the computer screen. Participants pressed the 'yes' key if the statement was true based on information that was either stated or implied in one of the passages they had read and the 'no' key if the statement was neither stated nor implied. Participants' responses and reaction times were recorded by the computer.

Inferring-New-Word-Meanings-From-Context Task

This was Daneman and Green's (1986) task in which readers had to infer the meaning of an unfamiliar word such as qualtagh from cues in the verbal context. The text did not provide direct synonyms for the word, but it did provide indirect cues. By piecing these clues together, readers could infer that qualtagh means 'the first person seen after leaving the house.' I predicted that this task would load heavily on the text inferencing component because of the requirement to piece together clues provided in the text. I predicted that the task would load heavily on the text memory component because of the sheer number of clues that readers would have to store (backtracking to previously read sentences was precluded in the paradigm). However, I recognized that the passages were sufficiently long and complex that the task would likely also draw on knowledge access and knowledge integration as well.

The stimulus materials consisted of 10 of the passages used by Daneman and Green (1986). Mean passage length was 224 words. Participants were told to read each passage carefully, using the cues in the context to determine the meaning of the target word. Each passage was presented one sentence at a time in the center of the computer screen. When the participant pressed the '+' key for the next sentence, the previous sentence was removed from the screen. Following presentation of the passage, participants wrote down as precise a definition for the target word as they could. They were warned that there was no single word synonym for any of the target words.

The quality of each definition was evaluated independently by two trained judges who assigned a score from 0 to 3. The interjudge reliability was .95; thus, an average of the two judges' ratings was used as a score. The 0 to 3 scoring system was designed to reflect the preciseness of meaning extracted and the degree of inference manifested beyond simple restatement of the contents of the cues themselves. To obtain a score of 2 or more, the definition had to exhibit acceptable levels of preciseness and inference.¹³

Memory-for-Details Task

In this task, participants read two 1500-word biographies and two 1000-word historical texts that were loaded with factual details. Immediately after reading each text, participants answered a series of multiple-choice questions designed to test memory for explicitly mentioned

¹³For example, for the word qualtagh, a score of 0 was assigned if the definition was completely off track (e.g., "one's personality or behavior"); a score of 1 was assigned if the definition was either too general or reflected a simple paraphrase of a contextual cue (e.g., "an encounter with a certain person will make or break your day"); a score of 2 was assigned if most but not all of the key elements of the definition were included (e.g., "the first person seen at the beginning of the day"); and a score of 3 was assigned if all the key elements of the definition were included (e.g., "the first person you encounter after leaving your place of residence"). Note that Daneman and Green (1986) used a 0 to 4 scoring system.

details. Although I predicted that the text memory component of the component processes task should be the best predictor of performance on these memory-loaded texts, I recognized that the texts were sufficiently lengthy and complex that they were likely to engage other components tapped by the task as well.

The biographical texts were about Julie Andrews and Carrie Fisher, and the historical texts were about the history of baroque art and the history of astronomy. They were typed, condensed versions of articles taken from The Current Biography Yearbook (Graham, 1994; Moritz, 1991) and Biographical Encyclopedia (Crystal, 1994). Each text was followed by 20 multiple-choice questions that tested memory for details mentioned explicitly in the text.

Participants were instructed to read each text carefully because they would be tested for details afterwards. They were given a maximum of 10 minutes to read each text. After reading the text, the 20 multiple-choice questions were presented one at a time on the computer screen. Presentation of questions was self-paced, but participants could not refer back to a previously answered question to help them answer a later one. All responses were recorded by the computer.

Results and Discussion

Table 8 presents the means and standard deviations on each component of the component processes task and on the global and specific tests of reading comprehension. Table 9 shows how the components correlated with each other and with the global test of reading comprehension ability. Because 22 of the participants in this experiment had also participated in Experiment 2, two sets of correlations are reported in Table 9, namely the correlations based on the 59 new participants and the correlations based on all 81 participants. Because the patterns

were so similar, the balance of the analyses are based on 81 participants.

The results of Experiment 3 replicated the Experiment findings with respect to the basic pattern of intercorrelations among the components of the new task and the ability of the components to predict performance on a global test of reading comprehension ability. As Table 9 shows, the two text-based components, text memory and text inferencing, were highly correlated with one another but, at best, only weakly correlated with the two knowledge access components. On the other hand, the three knowledge integration components were correlated with the text-based components and the knowledge access components. In addition, all components of the task were significantly correlated with performance on the Nelson-Denny test of reading comprehension, with the high-knowledge integration component being the best single predictor. A regression analysis was conducted on the reading comprehension data, using the same methodology as outlined in Experiment 2. The analysis showed that the component processes task accounted for an impressive 65% of the variance in reading comprehension performance, $F(4,76) = 34.83$, $MSe = 13.33$, $p < .001$, with the elementary components of text memory, low-knowledge access, and speed accounting for 59% of the variance, and the complex component of high-knowledge integration accounting for a further 6%.

Table 8

Means and Standard Deviations for Tasks in Experiment 3 (n = 81)

Test	<u>M</u>	<u>SD</u>	<u>Range</u>
<u>Component Processes Task^a</u>			
Text Memory	76.71	11.74	47.62 - 98.87
Text Inferencing	69.93	13.46	33.33 - 94.44
Low-Knowledge Integration	84.47	13.28	50.00 - 100
Medium-Knowledge Integration	80.83	13.38	47.22 - 100
High-Knowledge Integration	70.65	16.98	38.89 - 100
Low-Knowledge Access	90.74	7.27	69.40 - 100
High-Knowledge Access	86.73	8.64	66.67 - 100
Speed	4298.00	809.43	2870 - 6382
<u>Tests of Reading and Verbal Abilities</u>			
Reading Comprehension (max = 36)	24.12	5.99	11 - 34 ^b
Short-Stories Task			
Text Memory (max = 16)	15.08	1.08	12 - 16
True Text Inference (max = 8)	6.92	1.19	2 - 8
Verification Task			
Accuracy (max = 42)	38.53	2.31	30 - 42
Reaction Time	1630.90	369.52	969 - 3394
Bridging Inference Task			
Implicit Statements (max = 13)	11.42	1.49	6 - 13
Reaction Time Difference	118.21	377.22	-701 - 1959
Explicit Statements (max = 13)	12.17	1.27	8 - 13
Inferring-New-Word-Meanings-from-Context Task			
Definition Scores (max = 30)	12.24	5.58	1.5 - 25
Memory-for-Details Task			
Combined Texts (max = 80)	42.58	8.03	26 - 65
Biographies (max = 40)	24.44	4.88	13 - 37
Historical Texts (max = 40)	18.14	4.39	6 - 30

^aMeans, standard deviations, and ranges for the components are reported as percentages.

^bAccording to the Nelson-Denny norms, this range represents the 2nd to 97th percentiles (Brown et al., 1981) for university students.

Table 9

Correlations Among the Components of the Component Processes Task and Reading Comprehension in Experiment 3(n = 81)

	Text Inferencing	Low-Know Integration	Med-Know Integration	High-Know Integration	Low-Know Access	High-Know Access	Speed	Reading Comprehension
Text Memory	.72*** (.69)***	.62*** (.68)***	.68*** (.65)***	.61*** (.62)***	.33** (.36)**	.30** (.35)**	-.03 (.06)	.53**** (.55)***
Text Inferencing	- - -	.41*** (.46)***	.54*** (.52)***	.53*** (.52)***	.17 (.20)	.23* (.30)*	-.14 (-.06)	.40*** (.36)**
Low-Know Integration		- - -	.77*** (.83)***	.72*** (.80)***	.39*** (.41)**	.33** (.34)**	-.07 (-.03)	.53*** (.56)***
Medium-Know Integration			- - -	.76*** (.75)***	.39*** (.36)**	.30** (.26)	-.01 (-.03)	.51*** (.49)***
High-Know Integration				- - -	.39*** (.43)***	.41*** (.47)***	-.11 (-.12)	.61*** (.57)***
Low-Know Access					- - -	.56*** (.58)***	-.23* (-.25)	.55*** (.56)***
High-Know Access						- - -	-.32** (-.29)*	.41*** (.41)**
Speed							- - -	-.48*** (-.46)***

Note. Correlations for all 81 participants are displayed outside the parentheses and correlations for the new 59 participants are inside the parentheses.

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

So far, I have only considered how the component processes task predicts performance on a global test of reading comprehension ability. Next, I consider how the component processes task predicts performance on the five specific tests of reading comprehension ability. Table 10 shows how the components of the task correlated with the specific tests of reading comprehension. Table 11 shows the results for the stepwise regression analyses. The primary goal was to determine which of the components best predicted performance on each reading task. Consequently, for each dependent measure of interest, all seven components of the task plus the speed component were allowed to enter as predictors. Table 11 reports those predictor variables that contributed a significant amount of unique variance and the order in which they entered into the equation. Because I was also interested in determining whether the component processes task was as good or better than a standardized test like the Nelson-Denny at accounting for variance on these specific tests of reading comprehension, I conducted two regression analyses for each dependent measure. In the first regression analysis, I allowed the Nelson-Denny scores to enter the regression equation after all the components predicting unique variance were entered; the results of these analyses are reported in the lefthand panel (panel_i) of Table 11. In the second regression analysis, Nelson-Denny scores were entered into the regression model before any of the components; the results of these analyses are depicted in the righthand panel (panel_{ii}) of Table 11.

In general, the regression analyses confirmed my predictions concerning which component of the task would be the major contributor to performance on each of the specific tests of reading comprehension. The regression analyses also showed that the component processes task tended to be a much better predictor of performance on the specific reading tasks

than was the standardized test of reading comprehension. As the analyses in panel_i show, the Nelson-Denny test did not account for a significant proportion of the residual variance on any of the specific tests of reading comprehension, with the exception of the complex memory-for-details task. And even when the Nelson-Denny was entered into the regression equation first (panel_{ii}), it accounted for less unique variance than did the components when they were entered into the regression equation first (panel_i), again with the exception of the memory-for-details task.

Short-Stories Task

The regression analyses confirmed the predictions about the Oakhill (1982) short-stories task.¹⁴ As Table 11(a)_i shows, the text memory component of the task was the best predictor of readers' abilities to correctly verify test statements that required memory for information mentioned explicitly in the short stories, accounting for 33% of the variance. The only other significant contributor was the text inferencing component which accounted for an additional 4% of the variance. On the other hand, as Table 11(b)_i shows, only the text inferencing component of the task predicted performance on the statements that required readers to make an inference based on information mentioned in the text, accounting for 17% of the variance. These findings lend support to the contention that the text memory component of the task taps memory for facts mentioned explicitly in the text, whereas the text inferencing component of the task taps text-based inferences.

¹⁴Analyses of the short-stories task were based on 69 participants rather than 81. This was because I was required to make changes to the wording of the instructions for this task after I had run the first 12 participants on the task, and so the data for these 12 could not be included.

Table 10

Correlations Among the Components of the Component Processes Task and the Specific Reading Comprehension Tasks in Experiment 3

	Short Stories		Verification		Bridging Inferences			New Meaning from Context	Memory for Details		
	Text Memory	Text Inferencing	Accuracy	Reaction Time	Implicit Statements	Net Reaction Time	Explicit Statements		Combined	Biographies	Histories
Text Memory	.57***	.37**	.26*	-.12	.42***	-.23*	.47***	.44***	.49***	.40***	.45***
Text Inferencing	.56***	.42***	.23*	-.13	.24*	-.10	.35**	.40***	.35**	.32***	.32***
Low Know Integ	.36**	.25*	.30**	-.10	.46***	-.39***	.15	.36**	.48***	.36**	.48***
Medium Know Integ	.41***	.37**	.31**	-.11	.44***	-.18	.25*	.36**	.49***	.33**	.53***
High Know Integ	.43***	.30*	.23*	-.14	.36**	-.29*	.28*	.30**	.38***	.28*	.39***
Low Know Access	.06	.13	.42***	-.32**	.11	-.35**	.09	.35**	.32**	.26*	.30**
High Know Access	.03	-.01	.36**	-.36**	-.02	-.38***	.05	.38***	.37***	.29**	.35**
Speed	.05	-.03	-.33**	.56***	.12	.10	.07	-.42***	-.42***	-.41***	-.32**

* p < .05, ** p < .01, *** p < .001.

Table 11

Regression Analyses on Specific Reading Comprehension Tasks with Components of Component Processes Task as Predictors

	<u>R</u>	<u>R²</u>	<u>ΔR²</u>	<u>F</u>	<u>R</u>	<u>R²</u>	<u>ΔR²</u>	<u>F</u>
(a) Short Stories - text memory								
(i) Text Memory	.57	.33	.33	32.37	(ii) Reading Comprehension Test	.31	.09	6.93
Text Inferencing	.61	.37	.04	5.03	Text Memory	.57	.33	22.73
Reading Comprehension Test	.61	.37	.00	0.01	Text Inferencing	.61	.37	4.94
(b) Short Stories - text inferencing								
(i) Text Inferencing	.42	.17	.17	13.93	(ii) Reading Comprehension Test	.25	.06	4.59
Reading Comprehension Test	.43	.18	.01	0.81	Text Inferencing	.43	.18	9.53
(c) Verification Task - accuracy								
(i) Low-Knowledge Access	.42	.17	.17	16.25	(ii) Reading Comprehension Test	.40	.16	14.95
Speed	.48	.23	.06	5.51	Low-Knowledge Access	.47	.22	5.34
Reading Comprehension Test	.49	.24	.01	1.41	Speed	.49	.24	2.65

Table 11 continued

	R	R ²	ΔR ²	F	B	R ²	ΔR ²	F
<u>(d) Verification Task - reaction time</u>								
(i) Speed	.56	.32	.32	35.88	(ii) Reading Comprehension Test	.41	.16	15.31
Low-Knowledge Access	.60	.36	.04	4.33	Speed	.59	.34	20.59
Reading Comprehension Test	.60	.36	.00	.62	Low-Knowledge Access	.60	.36	1.90
<u>(e) Bridging Inference Task - accuracy for implicit passages</u>								
(i) Low-Knowledge Integration	.46	.21	.21	20.22	(ii) Reading Comprehension Test	.25	.06	5.24
Reading Comprehension Test	.46	.21	.00	0.00	Low-Knowledge Integration	.46	.21	13.85
<u>(f) Bridging Inference Task - net reaction time</u>								
(i) Low-Knowledge Integration	.39	.15	.15	13.72	(ii) Reading Comprehension Test	.20	.04	3.09
High-Knowledge Access	.47	.22	.07	6.69	Low-Knowledge Integration	.39	.15	10.12
Reading Comprehension Test	.48	.23	.01	0.94	High-Knowledge Access	.48	.23	7.59

Table 11 continued

	R	R ²	ΔR ²	F		R	R ²	ΔR ²	F
<u>(g) Bridging Inference Task - accuracy for explicit passages</u>									
(i) Text Memory	.47	.22	.22	21.46	(ii) Reading Comprehension Test	.14	.02	.02	1.45
Reading Comprehension Test	.49	.24	.02	2.18	Text Memory	.49	.24	.22	22.13
<u>(h) Inferring-New-Word-Meanings-from-Context Task</u>									
(i) Text Memory	.44	.19	.19	18.91	(ii) Reading Comprehension Test	.52	.27	.27	29.86
Speed	.60	.36	.17	20.26	Text Memory	.56	.31	.04	4.03
Reading Comprehension Test	.62	.38	.02	2.44	Speed	.62	.38	.07	8.59
<u>(l) Memory-for-Details Task - combined texts</u>									
(i) Text Memory	.49	.24	.24	24.55	(ii) Reading Comprehension Test	.67	.45	.45	63.37
Speed	.64	.41	.17	22.12	Text Memory	.69	.47	.02	3.50
Medium-Knowledge Integration	.68	.46	.05	7.18	Speed	.71	.50	.03	4.67
Reading Comprehension Test	.72	.52	.06	9.99	Medium-Knowledge Integration	.72	.52	.02	3.20

Table 11 continued

	R	R ²	ΔR^2	F	R	R ²	ΔR^2	F
(j) Memory-for-Details Task - biography texts								
(i) Speed	.41	.17	.17	16.19	(ii) Reading Comprehension Test	.60	.35	43.37
Text Memory	.57	.32	.15	17.13	Speed	.61	.38	2.62
Medium-Knowledge Integration	.57	.33	.01	.96	Text Memory	.63	.40	2.71
Reading Comprehension Test	.63	.40	.07	8.67	Medium-Knowledge Integration	.63	.40	.02
(k) Memory-for-Details Task - history texts								
(i) Medium-Knowledge Integration	.53	.28	.28	30.86	(ii) Reading Comprehension Test	.56	.31	35.64
Speed	.61	.38	.10	11.86	Medium-Knowledge Integration	.63	.39	10.46
Text Memory	.62	.39	.01	1.65	Speed	.64	.42	2.92
Reading Comprehension Test	.65	.39	.03	3.67	Text Memory	.65	.42	0.33

Note. ΔR^2 = change in R².

Verification Task

It was predicted that the knowledge access components (either low- or high-knowledge access) would be the best predictors of performance on the Smith et al. (1974) verification task. The correlations and regression analyses largely confirmed these predictions. As seen in Table 10, both low-knowledge access and high-knowledge access correlated significantly with performance on the verification task. As the regression analysis in Table 11(c_i) shows, when accuracy was the dependent measure on the verification task, the low-knowledge access component of the task was the best predictor of performance, accounting for 17% of the variance. The only other significant contributor was the speed component which accounted for a further 6% of the variance. When reaction time was the dependent measure, the knowledge access component was still a significant predictor; however, not surprisingly it was not as good a predictor as speed; see Table 11(d_i). Thus, the results for the verification task provide some support for the contention that the knowledge access components of the task tap the processes involved in accessing real-world knowledge from long-term memory.

Bridging Inference Task

I predicted that the knowledge integration component (either low-, medium-, or high-knowledge integration) would be the best predictor of both accuracy at verifying statements based on implicit passages and net reaction time (reaction time for implicit statements minus reaction time for explicit statements). As Table 11(e_i) and 11(f_i) show, the regression analyses confirmed both of these predictions.¹⁵ The low-knowledge integration component was the best

¹⁵Analyses of the bridging inference task were based on 79 participants rather than 81. The data for two participants could not be used because of technical difficulties with the computer program when running the task

predictor of accuracy at verifying implicit statements. In fact, it accounted for 21% of the variance, and no other component accounted for a significant proportion of the variance once low-knowledge integration was partialled out. This finding is consistent with that of Singer et al. (1992), who also found a significant correlation between Potts and Peterson's (1985) knowledge integration component and accuracy at making bridging inferences. However, Potts and Peterson's knowledge integration component accounted for only 5.3% of the variance in accuracy on Singer et al.'s (1992) near-bridging-inference task, which is considerably less than the 21% variance that was captured by the knowledge integration component of the modified task. The low-knowledge integration component was also the best predictor of the net difference in reaction time for implicit versus explicit statements, although now the high-knowledge access component also predicted a significant proportion of additional variance. These findings lend support to the contention that the knowledge integration component taps access to prior knowledge and the integration of this knowledge with text information. Note also that it was the text memory component and not the knowledge integration component that predicted accuracy on explicit statements; it accounted for 22% of the variance; see Table 11 (g). Again this makes sense given that verifying the explicit statements required memory for an explicitly stated fact rather than a knowledge-based inference.

An advantage of an individual-differences approach such as this one is that it can sometimes resolve apparent inconsistencies in the literature that arise from considering group means only. Indeed, the group reaction time data in this experiment failed to replicate Keenan and Kintsch's (1974) findings in that I observed a statistically significant 118 ms disadvantage for verifying implicit statements versus explicit statements, $t(78) = 2.79$, $p < .007$ (see Table 8),

whereas Keenan and Kintsch found no difference. Thus, the present group data suggested that individuals had not made the bridging inference at the time of reading the passage, whereas Keenan and Kintsch's data suggest that they had. However, this inconsistency can be reconciled if the reaction time data are examined as a function of individual readers' knowledge integration skills. An analysis of variance (ANOVA) on the present reaction time data with passage type (implicit vs. explicit) as one factor and knowledge integration skill (skilled vs. less-skilled on the low-knowledge integration component) as the other factor showed a highly significant Passage Type x Knowledge Integration Skill interaction, $F(1, 77) = 7.46$, $MSe = 65707$, $p < .01$. Readers who were less skilled at knowledge integration were 273 ms slower on the implicit statements than on the explicit statements, $t(26) = 2.81$, $p < .01$, suggesting that they had frequently failed to make the bridging inferences on-line. In contrast, readers who were skilled at knowledge integration showed the Keenan and Kintsch pattern of a nonsignificant 38 ms difference, $t(51) = 1.04$, $p > .30$. This finding suggests that readers differ in their tendency or ability to make bridging inferences, and this difference is captured by the knowledge integration component of the component processes task.

Inferring-New-Word-Meanings-from-Context Task

As predicted, the complexity of the passages resulted in all components of the component processes task being significant predictors of how well readers could infer new word meanings from text. However, also as predicted, text inferencing and text memory were particularly good predictors (see Table 10). Text memory was the slightly better predictor of the two, and because the two were highly correlated (.72), text inferencing did not contribute a significant proportion of the variance after the effects of text memory were statistically removed in a stepwise

regression analysis; see Table 11(h_i). Indeed, speed was the only other component to contribute significantly after the effects of text memory were statistically removed.

Memory-for-Details Task

As anticipated, the complexity of the biographical and historical passages resulted in significant correlations between all components of the component processes task and performance on the detail multiple-choice questions (see Table 10). A stepwise regression analysis on the combined data for the biographical and historical texts showed that text memory was the best predictor of performance on the detail questions (it accounted for 24% of variance), with speed and the medium-integration component also contributing significant unique variance; see Table 11(i). However, closer inspection of the data revealed slightly different patterns for the biographies and the historical texts. As Table 11(j_i) and 11(k_i) show, text memory was a better predictor than knowledge integration for the biographies, whereas knowledge integration was a better predictor than text memory for the historical texts. In retrospect, this is not surprising. The historical texts were more difficult than the biographical texts, and presumably required more access to and integration of prior knowledge with new text information than did the straightforward biographies.

Notice also that the memory-for-details task was the only task in which the Nelson-Denny reading comprehension test accounted for as much variance as did the components of the new task. Of all the specific tests used in Experiment 3, the memory-for-details one was the most like the Nelson-Denny in terms of passage length and complexity as well as multiple-choice testing format, and so it is not surprising that the Nelson-Denny test accounted for additional variance unaccounted for by the components of my new task.

Summary

In general, I think that the results of Experiment 3 are highly encouraging for at least three reasons. First, the results of Experiment 2 were replicated inasmuch as the new component processes task again accounted for a substantial proportion of the variance in performance on the Nelson-Denny test of reading comprehension ability. Second, this experiment showed that the predictive power of the new task was not confined to one measure of reading comprehension, namely, the Nelson-Denny, and one method of assessing comprehension, namely, a multiple-choice recognition method. Indeed, components of the task were good predictors of performance on a set of specific reading comprehension tasks that drew on a variety of text types (e.g., single sentences, short paragraphs, complex paragraphs), assessment methods (e.g., true-false judgments, definition generation, multiple choice), and dependent measures (accuracy, reaction time, net reaction time). And finally, Experiment 3 provided further construct validity for the separate components of the task by showing that each component was the best predictor of performance on a specific test of reading comprehension that was designed to load more heavily on that component.

Chapter 6: Experiment 4

In the final experiment, I pitted the component processes task against working memory span, another theoretically motivated measure that has been shown to be a good predictor of reading comprehension ability (e.g., Baddeley et al., 1985; Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Dixon et al. 1988; Masson & Miller, 1983). This experiment allowed me to investigate the relative predictive powers of the two tasks and the extent to which they make overlapping or independent contributions in accounting for individual differences in reading comprehension ability.

As my measure of working memory span, I used a variant of Daneman and Carpenter's (1980) reading span test which was designed to measure the combined processing and storage capacity of working memory during reading. In this version (see also Daneman & Hannon, 2001), participants read aloud increasingly longer sets of unrelated sentences (e.g., Torrential rains swept over the tiny deserted island; They ran until their lungs felt like they were going to burst; The starving hamburger bit into the juicy man.) and made judgments about the sensibility of each sentence. Then at the end of a set, they were required to recall the final word of each sentence in the set. (e.g., island; burst; man). Daneman and Carpenter (1980) found that measures of the combined processing and storage capacity of working memory were better predictors of reading comprehension performance than were measures of only storage capacity, such as the traditional digit span and word span measures. On average, Daneman and Carpenter's (1980) process plus storage measures of working memory capacity accounted for 31% of the variance in performance on a global test of comprehension (the VSAT), and 58% of

the variance in performance on specific tests of comprehension, such as retrieving facts and computing the antecedent referent for a pronoun. Since the original Daneman and Carpenter (1980) study with 41 participants, many researchers have used variants of reading span to investigate the association between working memory capacity and a variety of language comprehension tasks. A recent meta-analysis of this literature (Daneman & Merikle, 1996) supported most of Daneman and Carpenter's (1980) findings, although the working memory/comprehension correlations based on 6,179 participants were smaller in magnitude than those based on only 41 participants, with verbal process plus storage measures accounting for 17% of the variance on global tests and 27% on specific tests. According to the theory, working memory span is a good predictor of comprehension because readers who have less capacity to process and store information simultaneously in working memory are at a disadvantage when it comes to integrating successive ideas in a text because they have less capacity to keep the earlier read relevant information still active in working memory (Daneman & Carpenter, 1980, 1983).

The goals of this experiment were to compare the power of the new component processes task with that of a typical working memory span task at predicting global reading comprehension ability, and to investigate the extent to which the two tasks tapped common processes. One could argue that it is unfair to compare the predictive power of a four-component task with that of a single-component task such as working memory. However, researchers have argued that working memory span is such a good predictor of reading comprehension precisely because working memory is a resource shared by many of the component processes of comprehension (see Daneman, 1991; Just & Carpenter, 1987), and so in this sense, pitting the two tasks against each other does not seem unreasonable. But either way, the real theoretical interest was in

determining the extent to which the two measures tap similar or different component processes. I predicted that working memory span would be correlated with the text memory, text inferencing, and knowledge integration components because working memory seems to be important for the processes that require the maintenance and integration of successively encountered ideas in a text. On the other hand, there is no a priori reason for predicting that working memory measures should be related to the knowledge access component (but see Rosen & Engle, 1997) or to the speed component of the task (see Rankin, 1993). Consequently, I predicted that the component processes task would account for more of the variance on a global reading comprehension test than should the working memory span task because it captures variance associated with prior knowledge access and processing speed as well variance associated with text-based processes.

Method

Participants

The participants were 94 students from the University of Toronto community. All were fluent in English and had not participated in any of the previous experiments. In addition to being administered the component processes task and the Nelson-Denny test of reading comprehension, participants were administered a test of working memory capacity (Daneman & Carpenter, 1980) described below.

Working Memory Span Test

Participants read aloud a set of unrelated sentences, made a sensibility judgment for each one, and then at the end of the set, they recalled the last word of each sentence in the set.

Sentences 8 to 12 words in length, each ending with a different word, were presented one at a

time on the computer screen. After the participant responded 'yes' or 'no' to indicate whether or not the sentence made sense, the participant pressed the space bar for the next sentence. The procedure was repeated until a blue screen indicated that the trial was over, at which point the participant recalled the last word of each of the sentences in the set. Participants were allowed to recall the words in any order but were encouraged not to recall the last word in the set first. For example, in a two-sentence set, participants might read An eerie breeze suddenly chilled the warm, humid air. The umbrella grabbed its bat and stepped up to the plate. They would respond 'yes' after reading the first sentence, 'no' after reading the second sentence, and then recall 'air' and 'plate' when prompted by the blue screen. Sentences were arranged in five sets each of 2, 3, 4, 5, and 6 sentences. Participants were presented with increasingly longer sentence sets until all 100 sentences had been presented. Working memory span was the total number of sentence-final words out of 100 that the participant could recall.

Results and Discussion

Table 12 presents the means and standard deviations on each component of the component processes task and on the tests of global reading comprehension and working memory span. Table 13 shows how the components correlated with each other and with the reading comprehension and working memory span tests.

Once again, the earlier findings were replicated with respect to the basic pattern of intercorrelations among the components of the new task and the ability of each component to predict performance on a global test of reading comprehension (see Table 13). Of course, with each new set of participants, the pattern and magnitude of correlations on the new task do not remain identical. In Chapter 7, I report across-experiment analyses based on the data for all

participants who were administered the component processes task.

The main goals of Experiment 4 were to determine the extent to which the new component processes task and a working memory span task tapped similar processes, and to compare the powers of the two at predicting global reading comprehension performance. Table 13 shows that working memory span was significantly correlated with performance on the Nelson-Denny test of reading comprehension ($r = .46$). Working memory span was also significantly correlated with the text memory, text inferencing, and knowledge integration components of the component processes task (range = .36 to .48) but not with the knowledge access components or speed (range = -.03 to .21).¹⁶

¹⁶ A multiple regression analysis revealed that the component processes task accounts for 28% of the variance in performance on a test of working memory. Text memory and medium-knowledge integration accounted for 26% of the variance and the speed component accounted for about 2% of the variance.

Table 12

Means and Standard Deviations for Tasks Used in Experiment 4 (n = 94)

Test	<u>M</u>	<u>SD</u>	<u>Range</u>
<u>Component Processes</u> ^{Task}			
Text Memory	74.39	14.21	42.86 - 100
Text Inferencing	66.87	15.85	27.78 - 94.44
Low-Knowledge Integration	83.11	13.27	45.83 - 100
Medium-Knowledge Integration	80.85	12.91	47.22 - 100
High-Knowledge Integration	69.56	16.02	38.89 - 100
Low-Knowledge Access	91.76	6.12	69.44 - 100
High-Knowledge Access	88.25	8.38	58.33 - 100
Speed	4349.11	780.43	2811 - 6387
<u>Tests of Reading and Verbal Abilities:</u>			
Reading Comprehension (max = 36)	22.80	6.17	11 - 35 ^b
Working Memory Span (max = 100)	59.11	11.45	37 - 95

*Means, standard deviations, and ranges for the components are reported as percentages.

^bAccording to the Nelson-Denny norms, this range represents the 2nd to 99th percentiles (Brown et al., 1981) for university students.

Table 13
 Correlations Among Components of Component Processes Task, Reading Comprehension, and Working Memory Span
 in Experiment 4 (n = 94)

	Text Inference	Low Knowledge Integration	Medium Knowledge Integration	High Knowledge Integration	Low Knowledge Access	High Knowledge Access	Speed	Reading Comprehension	Working Memory
Text Memory	.77***	.62***	.73***	.77***	.17	.33**	.24**	.46***	.48***
Text Inference	---	.50***	.57***	.63***	.17	.38***	.17	.48***	.42***
Low-Knowledge Integration		---	.78***	.61***	.30**	.30**	.22*	.36***	.36***
Medium-Knowledge Integration			---	.69***	.35***	.36***	.21*	.46***	.46***
High-Knowledge Integration				---	.26*	.39***	.07	.58***	.44***
Low--Knowledge Access					---	.22*	-.10	.26*	.21*
High-Knowledge Access						---	.02	.30**	.19
Speed							---	-.28**	-.03
Reading Comprehension								---	.46***

* p < .05, ** p < .01, *** p < .001.

To determine the extent to which the new component processes task and working memory span made overlapping or independent contributions in their prediction of reading comprehension, two stepwise regression analyses were conducted, one in which working memory span was entered as the first predictor and a second one in which it was entered after the components of the component processes task were partialled out. As Table 14(a) shows, when entered first, working memory span accounted for 21% of the variance in reading comprehension performance, a finding that is consistent with the literature showing a reliable correlation between measures of working memory span and comprehension (see Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Turner & Engle, 1989). However, note also that the text inferencing, speed, and high-knowledge integration components accounted for a further 29% of the variance in reading after the effects of working memory span were removed. On the other hand, as Table 14(b) shows, when working memory span was entered into the regression equation after the 47% of variance accounted for by the component processes was partialled out, working memory accounted for only an additional 3% of unique variance. Thus, it would appear that the new component processes task is accounting for most of the variance in reading comprehension that is tapped by a typical test of the combined processing and storage capacity of working memory. It would also appear that the working memory span test shares most variance in common with the text inferencing component of the component processes task. This is because text inferencing is the component whose predictive power is most reduced by entering working memory span as the first predictor. These findings should be of theoretical interest to the increasing number of researchers interested in determining what working memory span measures such as reading span are really measuring (see Miyake & Shah, 1999, for a recent

review).

Summary

The results of Experiment 4 provide further support for the reliability of the component processes task in predicting performance on a global test of reading comprehension. In addition, the results show that the component processes task accounts for most of the variance in reading comprehension that is accounted for by a typical measure of working memory capacity; and it accounts for variance not accounted for by working memory, such as variance associated with access to prior knowledge, and speed of reading and responding.

Table 14

Regression Analyses on Reading Comprehension Scores in Experiment 4 (n = 94)
with Working Memory Span and Component Processes as Predictors

	Variable	R	R ²	ΔR^2	F
(a) Working memory as first predictor of reading comprehension					
1.	Working memory	.46	.21	.21	24.43
2.	Text inferencing	.56	.31	.10	13.07
3.	Speed	.65	.42	.11	16.62
4.	High-knowledge integration	.71	.50	.08	14.65
(b) Working memory as predictor of reading comprehension after variance accounted for by the component processes has been partialled out					
1.	Text inferencing	.48	.23	.23	27.23
2.	Speed	.60	.36	.13	18.21
3.	High-knowledge integration	.69	.47	.11	19.38
4.	Working memory	.71	.50	.03	5.02

Note. ΔR^2 = change in R².

Chapter 7: Analysis of Pooled Data from the Previous Experiments

Given that there is now data on the component processes task for 222 participants (69 in Experiment 2; 59 in Experiment 3; 94 in Experiment 4), it might be useful to examine the pooled data. In this chapter I report the correlations, a confirmatory factor analysis, and regression analyses based on the entire set of participants.

Table 15 provide the correlations based on the entire set of participants. Note that the pattern of correlations for all 222 participants is remarkably similar to the pattern for the original subset of 69 who were tested on the task in Experiment 2 (Chapter 4). The pattern shows that the two text-based components, text memory and text inferencing, were highly correlated with one another, and both were at best weakly correlated with the knowledge access components. On the other hand, all three knowledge-integration components were significantly correlated with the two text-based components, and they were also correlated with the knowledge access components. Taken together, these findings suggest that there are two elementary sets of processes, text-based processes (measured by the text memory and text inferencing components), and knowledge access processes (measured by the low- and high-knowledge access components), that are at best weakly correlated with one another, and there is a more complex set of processes, knowledge integration processes (measured by the low-, medium-, and high-knowledge integration components) that are correlated with and dependent upon both the text-based processes and knowledge access processes.

Table 15

Correlations Among the Components of the Component Processes Task and Reading Comprehension across Experiments 2, 3, and 4 (n = 222)

	Text Inference	Low-Knowledge Integration	Medium-Knowledge Integration	High-Knowledge Integration	Low-Knowledge Access	High-Knowledge Access	Speed	Reading Comprehension
Text Memory	.77***	.62***	.69***	.70***	.24***	.28***	.11	.47***
Text Inference	---	.50***	.55***	.58***	.20**	.31***	.05	.44***
Low-Knowledge Integration		---	.77***	.65***	.37***	.31***	.08	.42***
Medium-Knowledge Integration			---	.71***	.39***	.30***	.08	.45***
High-Knowledge Integration				---	.34***	.35***	-.03	.58***
Low-Knowledge Access					---	.36***	-.17	.39***
High-Knowledge Access						---	-.12	.34***
Speed							---	-.37***

Note. Cronbach alpha reliability estimates for all components ranged from .83 to .86.

* p < .05, ** p < .01, *** p < .001.

To support this interpretation of the component processes task, a confirmatory factor analysis (CFA) on the pooled data was conducted. There were two steps to the analysis, as recommended by Jöreskog and Sörbom (1993). The first step was a measurement model that confirmed that the component processes task does indeed consist of a text processing latent variable that is measured by the observed variables, text memory and text inferencing; a knowledge access latent variable that is measured by the observed variables, low-knowledge access and high-knowledge access; and a knowledge integration latent variable that is measured by the observed variables, low-knowledge integration, medium-knowledge integration, and high-knowledge integration. The second step was a structural model of the component processes task that confirmed the interrelation among the latent variables; in particular, it confirmed that there are two elementary latent variables, text processing and knowledge access, that are both necessary for the more complex latent variable, knowledge integration. For both steps of the CFA, I used LISREL 8, a computer package designed for structural equation modeling (Jöreskog & Sörbom, 1993). Table 16 contains the results of the two steps of the CFA.

The measurement model fit the data well. As Table 16 shows, the results of the CFA for the measurement model confirmed that the text memory and text inferencing components of the component processes task do indeed measure one latent variable which I interpret as text processing, that the low- and high-knowledge access components do indeed measure a second latent variable which I interpret as knowledge access, and that the low-, medium-, and high-knowledge integration components do indeed measure a third latent variable which I interpret as knowledge integration.

Table 16

Results of CFA Analyses for Measurement and Structural Models (n = 222)(a) Results for Measurement Model

	<u>R</u>	<u>R²</u>	<u>z-score</u>
<u>Text Processing</u>			
Text Memory	.98	.95	18.20*
Text Inferencing	.79	.62	13.48*
<u>Knowledge Access</u>			
Low-real Knowledge Access	.64	.41	7.32*
High-real Knowledge Access	.56	.31	6.71*
<u>Knowledge Integration</u>			
Low-knowledge Integration	.77	.59	8.15*
Medium-knowledge Integration	.84	.71	7.00*
High-knowledge Integration	.85	.72	6.96*

(b) Fit Statistics for Structural Model

<u>df</u>	<u>χ^2</u>	<u>χ^2 p(exact)</u>	<u>χ^2 p(close)</u>	<u>GFI</u>	<u>AGFI</u>	<u>CFI</u>	<u>RMSEA</u>
10	11.59	0.31	0.70	0.99	0.96	1.00	0.027

Note. p(exact) and p(close) reflect the significance levels for the χ^2 (exact) and χ^2 (close) tests. CFA = confirmatory factor analysis; GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; CFI = comparative fit index; RMSEA = root-mean-square error of approximation. *p < .05.

For the structural model, six different fit statistics were conducted according to the recommendations of Hoyle and Panter (1995). These fit statistics were the traditional chi-square test of 'exact' model fit, the chi-square test of 'close' model fit (Browne & Cudeck, 1993), the goodness-of-fit index (GFI), the adjusted GFI (AGFI; Jöreskog & Sörbom, 1981), the root-mean-square error of approximation (RMSEA; Steiger & Lind, 1980), and the comparative fit index (CFI; Bentler, 1989). It is important to note that for the chi-square tests, the hypothesis being tested assumes either an exact model fit for the data or a model fit that is acceptable, and so a good-fitting model is indicated by nonsignificant results from these tests (i.e., $p > .05$); see also Engle, Tuholski, Laughlin, and Conway (1999). For the GFI, AGFI, and CFI indices, I followed the general guideline that good-fitting models are indicated by a value of .90 or more, and for interpreting the RMSEA statistic, I followed the recommendation that values of .05 or less indicate a good-fitting model (Browne & Cudeck, 1993). As Table 16 indicates, the results of all six statistical tests suggested that there was an acceptable model fit for the structural model of two elementary latent variables, text processing and knowledge access, both of which influence a more complex latent variable, knowledge integration.

Figure 1 shows the parameter estimates for the measurement and structural equation models. In this figure, the ellipses represent the latent variables (e.g., text processing) and the rectangles represent the observed measures/variables (e.g., text memory, text inferencing). Paths leading from one variable to another variable indicate direction of influence. Numbers on paths are standardized path coefficients that indicate the degree of influence. As Figure 1 shows, the latent variable text processing influences the observed variables, text memory and text inferencing; the latent variable knowledge integration influences the observed variables, low-,

medium-, and high-knowledge integration; the latent variable knowledge access influences the observed variables low- and high-knowledge access; and the latent variables text processing and knowledge access both influence the latent variable, knowledge integration.

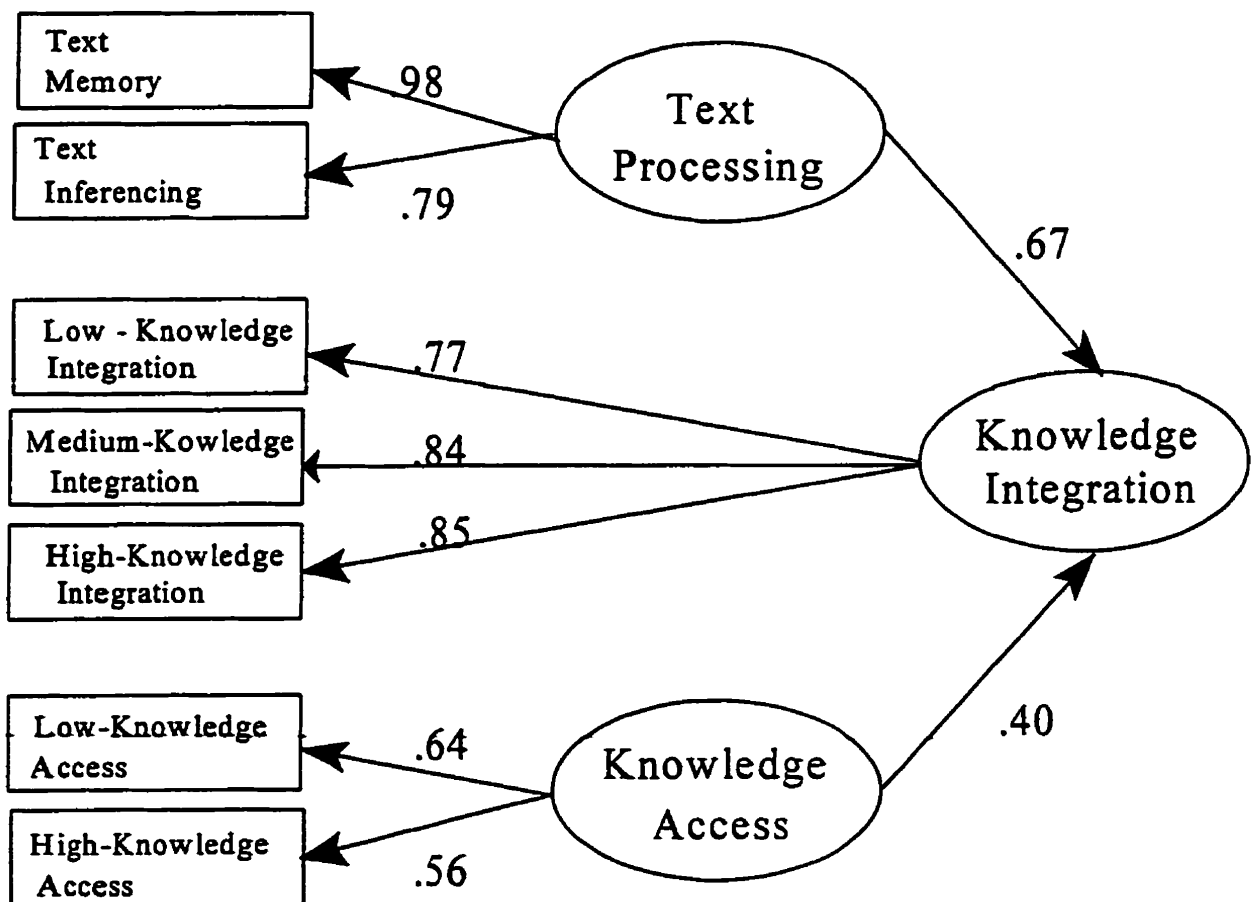


Figure 1. Path Diagram Showing the Measurement Model and the Structural Equation Model for the Component Processes Task (n=222). All paths are significant at the .05 level.

So far, only the relationships among the components of the component processes task have been considered. The final step was to investigate how well the components accounted for variance on the Nelson-Denny reading comprehension test. To do this, two regression analyses were conducted on the data pooled over all 222 participants. These analyses are presented in Table 17, and again they are similar to those reported for the first 69 participants in Experiment 2 (Chapter 4).

For the regression analysis reported in Table 17(a), the procedure was the same as the one used in Experiment 2. That is, the elementary components of text memory, text inferencing, and knowledge access were allowed to enter the equation before the more complex knowledge-integration components. As Table 17(a) shows, the component processes task accounted for 52% of the variance in reading performance for the 222 participants, with the elementary components of text memory, speed, low-knowledge access, and text inferencing accounting for 46% of the variance, and the complex component of high-knowledge integration accounting for a further 6%. Given the large sample size, I feel confident that these results are stable and reliable.

Table 17(b) shows the results of a regression analysis using the same procedure as was used for the specific reading tests in Experiment 3. That is, all eight components of the task were allowed to enter the regression analysis freely, and only those components that contributed significantly are reported. This analysis allowed a comparison between the pattern of predictor variables for the global reading comprehension test with the patterns for the specific reading tasks used in Experiment 3 (Chapter 5). As seen in Table 17(b), high-knowledge integration was the best single predictor of reading comprehension for the 222 participants as it was for the original subset of 69 participants, and it accounted for 34% of the variance in reading

comprehension performance on the Nelson-Denny. The pattern in Table 16(b) is most similar to that for the historical texts in Experiment 3 (see Table 11[k₁]), except that for the historical texts, the medium-knowledge integration component rather than the high-knowledge integration component was the best predictor of comprehension performance.

Table 17

Regression Analyses on Reading Comprehension Scores for all 222 Participants
with Components of Component Processes Task as Predictors

Variable	<u>R</u>	<u>R²</u>	<u>Δ R²</u>	<u>F</u>
(a) Reading comprehension				
1. Text memory	.47	.22	.22	63.48
2. Speed	.64	.41	.18	67.53
3. Low-knowledge access	.67	.45	.04	15.82
4. Text inferencing	.68	.46	.01	3.67
5. High-knowledge integration	.72	.52	.06	27.10
(b) Reading comprehension				
1. High-knowledge integration	.58	.34	.34	113.05
2. Speed	.68	.47	.13	51.95
3. Text memory	.70	.49	.02	11.16
4. Low-knowledge access	.72	.51	.02	8.58
5. Text inferencing	.72	.52	.01	2.40

Note. Δ R² = change in R².

Chapter 8: General Summary and Conclusions

The goal of this research was to develop a new measure of individual differences in reading ability that is theoretically motivated, easy to administer, and has high predictive power. It appears that the component processes task fits the bill. The task measures four components—the ability to access prior knowledge from long-term memory, the ability to integrate accessed knowledge with text information, the ability to make inferences based on information provided in the text, and the ability to recall new text information from long-term memory. These four components plus a speed component account for a substantial proportion of the variance in the performance of university students on a global standardized test of reading comprehension performance (Experiments 2, 3, and 4). Indeed, the component processes task accounts for more variance in reading ability than do single-component tasks such as vocabulary knowledge (Experiment 2), and single-resource tasks such as working memory span (Experiment 4). Nevertheless, the task is almost as quick to administer as a typical working memory span task (takes approximately 30 minutes as compared to 20 minutes for the working memory span task). All in all, the component processes task is a powerful and practical research tool for measuring individual differences in reading comprehension ability, and it provides theoretical insights concerning the nature of the components underlying various reading comprehension tasks.

Skeptics might argue that there is nothing particularly surprising or informative about the finding that performance on the component processes task correlates well with performance on a standardized reading comprehension task like the Nelson Denny. After all, the task requires

people to read and comprehend short paragraphs and to answer questions about them, and the Nelson-Denny requires people to read and comprehend somewhat longer passages and to answer questions about them. From this perspective, I am the first to agree that the results are not surprising. In this thesis, I have claimed that the component processes task measures four component processes of reading comprehension, all four of which have been proposed (in one form or another) by reading researchers in the past, and I have shown that the measure of these four components did indeed correlate with a full-blown reading comprehension task. In that sense, the findings are not surprising. But to construe the findings as not theoretically informative is to misconstrue the value of a componential analysis of reading comprehension. I believe that the results are informative in a number of ways. Although this is not the first research to have considered text memory, text inferencing, knowledge access, and knowledge integration to be important components underlying reading comprehension and important sources of individual differences in reading comprehension, I believe that this is the first research to provide reliable and valid ways to measure them. This individual-differences approach has enabled me to provide new evidence concerning just how much variance in performance on a standardized reading comprehension test can be attributed to each of these four components, as well as the extent to which they make overlapping or independent contributions. These are all valuable theoretical contributions to the understanding of the complex construct of reading comprehension. So far, I have shown how the task predicts the reading performance of a fairly circumscribed population of university students. It is quite likely that the components that contribute to reading comprehension ability among university students may differ from the components that contribute to reading comprehension variability in younger readers, older

readers, and less-skilled readers. I anticipate that the task can be adapted to studying other populations of readers as well.

There are also several other theoretical contributions of this research worth mentioning. For starters, consider a rather interesting result that was mentioned but not emphasized so far. In Experiment 2 (Chapter 4), it was observed that vocabulary knowledge accounts for little unique variance in performance on a global test of reading ability when the component processes of the task are partialled out. This finding is theoretically interesting because it seems to suggest that the bulk of the shared variance between tests of reading comprehension and tests of vocabulary may be due to shared cognitive and inferential processes and not vocabulary knowledge per se (see also Daneman, 1988; Sternberg & Powell, 1983). Also of theoretical interest is the finding in Experiments 2 (Chapter 4), 3 (Chapter 5), and 4 (Chapter 6) that skill at accessing prior knowledge accounts for unique variance in reading comprehension ability. In other words, it is not simply how much relevant prior knowledge the reader has; after all, care was taken to ensure that all readers had the prior knowledge that an ostrich is larger than a robin. Equally important to reading success is the ability to have ready access to that long-term memory knowledge (see also Ericsson & Kintsch, 1995). Of theoretical interest to working memory researchers was the finding in Experiment 4 that working memory span is correlated with the text-based processing aspects of reading comprehension rather than access to prior knowledge, or speed. This finding supports Daneman and Carpenter's (1983) claim that working memory is implicated in the processes that require the maintenance and integration of successively encountered ideas in a text. And finally, in Experiments 2, 3, and 4 it was consistently found that both text-based components and knowledge-based components are important contributors to reading

comprehension success. This finding corroborates Gernsbacher's (1990) view of comprehension as the building of a mental structure based on information in the text and information in long-term memory.

The potential applications of the task are many. So far it has been shown how the various components of the component processes task can be used to predict performance on a typical test of global reading comprehension ability, namely the Nelson-Denny, as well as on a range of specific tests, each of which draws more heavily on one particular component of the task. But this research has only just scratched the surface. I believe that the component processes task can be used to measure and understand performance on a whole range of cognitive and linguistic tasks. For example, although not emphasized in this thesis, it was observed that the components of the task had the power to predict 37% of the variance on several verbal reasoning tasks (Experiment 2). In addition, preliminary results show that the component processes task can be used to measure and understand individual differences in performance on a variety of tasks that are of considerable theoretical interest to cognitive researchers. For example, the task is useful for measuring and understanding individual differences in people's susceptibility to semantic illusions (e.g., Erickson & Mattson, 1981; Reder & Kusbit, 1991), as when they respond "two" to the question "How many animals of each kind did Moses take on the ark?," even though they know that it was Noah, and not Moses, who took animals on the ark (Hannon & Daneman, 2001). These studies have shown that individuals who are less skilled at accessing and reasoning about real world knowledge (as measured by the high-knowledge access component) succumb to more semantic illusions than do individuals who are skilled at knowledge access, and this is particularly evident when the impostor word is highly related (Moses) to the intended target

(Noah) than when it is less highly related (Adam). The task is also proving useful for measuring and understanding individual differences in performance on a verbal fluency task (e.g., Rosen & Engle, 1997; Troyer, Moscovitch, & Winocur, 1997), in which individuals are given a minute or two to generate as many exemplars as they can for a given category (e.g. animals). I anticipate that the component processes task can be applied toward measuring and understanding the higher order cognitive processes underlying a whole host of tasks that depend on memory, inferencing, knowledge access, and knowledge integration.

Finally, I recommend that researchers include this individual-differences task even if the primary goal of their research is not the investigation of individual differences per se. One of the problems with developing models based on group data is that the group data frequently mask interesting qualitative differences among individuals in the way the task is performed. This may lead to spurious models of how the task is performed. Moreover, it may be the cause of inconsistent findings in the literature whenever there are across-study differences in the skill level of participants. For example, we saw how the inclusion of an independent measure of knowledge-integration skill allowed me to uncover a plausible reason for why Keenan and Kintsch's (1974) group data suggest that readers were making bridging inferences, whereas the group data in Experiment 3 (Chapter 5) suggest that readers were not. It is quite likely that many other inconsistencies in the literature might be understood if an appropriate measure of individual differences is included in the studies.

References

- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.) Handbook of reading research (Vol 1. pp. 225-253). New York: Longman.
- Arenberg, D. (1968). The concept problem solving in young and old adults. Journal of Gerontology, 23, 279-282.
- Baddeley, A. D., Logie, R., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. Journal of Memory and Language, 24, 119-131.
- Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. Journal of Educational Psychology, 74(4), 506-521.
- Bell, L. C., & Perfetti, C. A. (1994). Reading skill: Some adult comparisons. Journal of Educational Psychology, 86, 244 - 255.
- Bentler, P. B. (1989). EQS structural equations program manual. Los Angeles: BMDP Statistical Software.
- Bloom, B. S. (1976). Human characteristics and school learning. New York: McGraw-Hill.
- Brown, J. I., Bennett, J. M., & Hanna, G. (1981). Nelson-Denny Reading Test. Chicago: Riverside.
- Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In K. A. Bollen and J. S. Long (Eds.) Testing structure equation models (pp. 136-142). Newbury Park, CA: Sage.

Carr, T. H. (1981). Building theories of reading ability: On the relation between individual differences in cognitive skills and reading comprehension. Cognition, *9*, 73-114.

Carroll, J. B. (1976). Psychometric tests as cognitive tasks: A new "structure of intellect." In L. B. Resnick (Ed.), The nature of intelligence (pp. 27 - 56). Hillsdale, NJ: Lawrence Erlbaum.

Carroll, J. B. (1978). How shall we study individual differences in cognitive abilities?—Methodological and theoretical perspectives. Intelligence, *2*, 87-115.

Chabot, R. J., Zehr, H. D., Prinzo, O. V., & Petros, T. V. (1984). The speed of word recognition subprocesses and reading achievement in college students. Reading Research Quarterly, *19*, 147-161.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.) The Psychology of Learning and Motivation (Vol 16, pp. 1-58). New York: Academic Press.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.) Cognitive Skills and their Acquisition (pp. 141-189). Hillsdale, NJ: Erlbaum.

Chiang, A., & Atkinson, R. C. (1976). Individual differences and interrelationships among a select set of cognitive skills. Memory and Cognition, *4*, 661-672.

Crystal, D. (Ed.). (1994). Biographical encyclopedia. Cambridge University Press.

Daneman, M. (1982). The measurement of reading comprehension: How not to trade construct validity for predictive power. Intelligence, *6*, 331-334.

Daneman, M. (1988). Word knowledge and reading skill. In M. Daneman, G. E. MacKinnon, & T. G. Waller (Eds.), Reading research: Advances in theory and practice (Vol. 6, pp. 145-175). New York: Academic Press.

Daneman, M. (1991). Individual differences in reading skill. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), Handbook of reading research (Vol. 2, pp. 512-538). New York: Longman.

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

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 561-584.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. Journal of Memory and Language, 25, 1-18.

Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. Journal of Experimental Psychology: General 130, 208-233.

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

Daneman, M., & Tardif, T (1987). Working memory and reading skill re-examined. In M. Coltheart (Ed.), Attention and performance XII (pp. 491-508). London: Erlbaum.

Davis, F. B. (1944). Fundamental factors of comprehension in reading. Psychometrika, 9(3), 185-197.

Davis, F. B. (1968). Research in comprehension in reading. Reading Research Quarterly, 3, 499 - 545.

Dixon, P., LeFevre, J., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. Journal of Educational Psychology, 80, 465-472.

Educational Testing Service for the Graduate Record Examinations Board(1992). Practicing to take the GRE: General Test - No. 9. Princeton, NJ: Warner Books

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 972 - 992.

Engle, R. W., Nations, J. K., & Cantor, J. (1990). Is "working memory capacity" just another name for word knowledge? Journal of Educational Psychology, 82, 799-804.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. Journal of Experimental Psychology: General, 128, 309-331.

Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. Journal of Verbal Learning and Verbal Behavior, 20, 540-551.

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

Frederiksen, J. R. (1982). A componential theory of reading skills and their interactions. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1, pp. 125-180). Hillsdale, NJ: Erlbaum.

Garner, R. (1980). Monitoring of understanding: An investigation of good and poor readers' awareness of induced miscomprehension of text. Journal of Reading Behavior, 12, 55-64.

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

Graham, J. (Ed.). (1994). Current biography yearbook. New York: H. W. Wilson Company.

Guilford J. P. (1981). Higher-order structure-of-intellect abilities. Multivariate Behaviour Research, 16(4), 411-435.

Haenggi, D., & Perfetti, C. A. (1994). Processing components of college-level reading comprehension. Discourse Processes, 17, 83-104.

Hannon, B., & Daneman, M. (1998). Facilitating knowledge-based inferences in less-skilled readers. Contemporary Educational Psychology, 23, 149-172.

Hannon, B., & Daneman, M. (2001). Susceptibility to semantic illusions: An individual-differences perspective. Memory & Cognition, 29 (3), 449-461.

Hayslip, B., Jr., & Sterns, H. L. (1979). Age differences in relationships between crystallized and fluid intelligences and problem solving. Journal of Gerontology, 34, 404-414.

Hogaboam, T. W., & Perfetti, C. A. (1978). Reading skill and the role of verbal experience in decoding. Journal of Educational Psychology, 70, 717-729.

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Panter (Ed.) Structural equation modeling: Concepts, issues, and applications. Thousand Oaks, CA: Sage.

Hunt, E. (1978). Mechanics of verbal abilities. Psychological Review, 85, 109-130.

Hunt, E. B. (1980). Intelligence as an information-processing concept. British Journal of Psychology, 71, 449-474.

Hunt, E., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? Cognitive Psychology, 7, 194-227.

Hunt, L. C. (1947). Can we measure specific factors associated with reading comprehension? Journal of Educational Research, 51, 161-172.

Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. Journal of Experimental Psychology: General, 108, 151-181.

Jöreskog, K., & Sörbom, D. (1981). LISREL V: Analysis of linear structural relationships by the method of maximum likelihood. Chicago: National Education Resources.

Jöreskog, K., & Sörbom, D. (1993). LISREL 8: Structural equation modeling with the SIMPLIS command language. Hillsdale, N.J.: Lawrence Erlbaum Associates.

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

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 99, 122 - 149.

Keenan, J. M., & Kintsch, W. (1974). Memory for information inferred during reading. In W. Kintsch (Ed.), The representation of meaning in memory (pp. 153-178). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kyllonen, P. C. (1993). Aptitude testing inspired by information processing: A test of the four-sources model. Journal of General Psychology, 120, 375-405.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity? Intelligence, *14*, 389-433.

Long, D. L., Oppy, B. J., & Seely, M. R. (1994). Individual differences in the time course of inferential processing. Journal of Experimental Psychology: Learning, Memory and Cognition, *20*, 1456-1470.

Lorch, R. F., Lorch, E. P., & Morgan, A. M. (1987). Task effects and individual differences in on-line processing of the topic structure of a text. Discourse Processes, *10*, 63-80.

Masson, M. E., & Miller, J. A. (1983). Working memory and individual differences in comprehension and memory of text. Journal of Educational Psychology, *75*, 314-318.

Miyake, A., & Shah, P. (1999). (Eds.) Models of working memory: Mechanisms of active maintenance and executive control. Cambridge: Cambridge University Press.

Moritz, C. (Ed.). (1991). Current Biography Yearbook. New York: H. W. Wilson Company.

National Institute for Testing and Evaluation. (1991). The inter-university psychometric entrance examination. Jerusalem, Israel.

Neisser, U. (1983). Components of intelligence or steps in routine procedures? Cognition, *15*, 189-197.

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

Oakhill, J., & Yuill, N. (1986). Pronoun resolution in skilled and less-skilled comprehenders: Effects of memory load and inferential complexity. Language and Speech, *29*, 25-37.

Olson, R. K., Kliegl, R., & Davidson, B. J. (1983). Dyslexia and normal readers' eye movements. Journal of Experimental Psychology: Human Perception and Performance, 9, 816-825.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. Cognition and Instruction, 1, 117-175.

Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. Journal of Memory and Language, 24, 59-88.

Pavlidis, G. T. (1981). Sequencing, eye movements, and the early objective diagnosis of dyslexia. In G. Pavlidis & T. Miles (Eds.), *Dyslexia research and its applications to education*. Winchester: Wiley.

Perfetti, C. A. (1976). Language comprehension the deverbalization of intelligence. In L.S. Resnick (Ed.), The nature of intelligence (pp. 283 - 292). Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (1985). Reading ability. New York: Oxford University Press.

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter, (Eds.), Cognitive processes in comprehension (pp. 141 - 183). Hillsdale, NJ: Lawrence Erlbaum Associates.

Peterson, L., & Peterson, M. J. (1959). Short term retention of individual verbal items. Journal of Experimental Psychology, 58, 193-198.

Posner, H. I., & Mitchell, R. F. (1967). Chronometric analysis of classification, Psychological Review, 74, 392-409.

Potts, G. R. (1977). Integrating new and old information. Journal of Verbal Learning and Verbal Behavior, 16, 305-320.

Potts, G. R., Keller, R.A., & Rooley, C. J. (1981). Factors affecting the use of world knowledge to complete a linear order. Journal of Experimental Psychology: Human Learning and Memory, 7(4), 254-268.

Potts, G. R., & Peterson, S. B. (1985). Incorporation versus compartmentalization in memory for discourse. Journal of Memory and Language, 24, 107-118.

Pressley, M. (1994). Embracing the complexity of individual differences in cognition: studying good information processing and how it might develop. Learning and Individual Differences, 6(3), 259-284.

Pressley, M., & Brewster, M. E. (1990). Imaginal elaboration of illustrations to facilitate fact learning: Creating memories of Prince Edward Island. Applied Cognitive Psychology 4(5), 359-369.

Rankin, J. L. (1993). Information-processing differences of college-age readers differing in reading comprehension and speed. Journal of Reading Behavior, 25, 261-278.

Rayner, K., & Duffy, S. A. (1988). On-line comprehension processes and eye movements during reading. In M. Daneman, G. E. MacKinnon, and T. G. Waller (Eds.), Reading research: Advances in theory and practice (Vol. 6, pp. 13-56). New York: Academic Press.

Reder, L. M., & Kusbit, G. W. (1991). Locus of the Moses illusion: Imperfect encoding, retrieval, or match? Journal of Memory and Language, 30, 385-406.

Robinson, F. P., & McCollum, F. H. (1934). Reading rate and comprehension accuracy as determinants of reading test scores. Journal of Educational Psychology, 25, 154-157.

Rosch, E. (1975). Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 104, 192-233.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. Journal of Experimental Psychology: General, 126, 211-227.

Samuels, S. J., LaBerge, D., & Bremer, C. D. (1978). Units of word recognition: Evidence for developmental changes. Journal of Verbal Learning and Verbal Behavior, 17, 715-720.

Schneider, W., Korket, J., & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. Journal of Educational Psychology, 81, 306-312.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. Journal of Experimental Psychology: General, 125, 4-27.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171, 701-703.

Shoben E. J. (1992). Semantic memory. In L. R. Squire (Ed.), Encyclopedia of learning and memory (pp. 581-585). New York: Macmillan.

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

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

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. Psychological Review, 81, 214-241.

Snowling, M. J. (1980). The development of grapheme-phoneme correspondence in normal and dyslexic readers. Journal of Experimental Child Psychology, 29, 294-305.

Spearman, C. (1923). The nature of "intelligence" and the principles of cognition. London: Macmillan.

Spilich, G. H., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high- and low-domain knowledge. Journal of Verbal Learning and Verbal Behaviour, 18, 275 - 290.

Spiro, R. J. (1980). Constructive processes in prose comprehension and recall. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.). Theoretical issues in reading comprehension (pp. 245 - 278). Hillsdale, NJ: Erlbaum.

Stanley, G., Smith, G., & Howell, E. (1983). Eye movements and sequential tracking in dyslexic and control children. British Journal of Psychology, 74, 181-187.

Stanovich, K. E., Cunningham, A., & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. Reading Research Quarterly, 19, 278-303.

Steiger, J. H., & Lind, J. C. (1980, May). Statistically based tests for the number of factors. Paper presented at the annual meeting of the Psychometric Society, Iowa City. IA.

Sternberg, R. J. (1977). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1979). The nature of mental abilities. American Psychologist, 34, 214-230.

Sternberg, R. J. (1980). Sketch of a componential subtheory of human intelligence. Behavioral and Brain Sciences, 3, 573-614.

Sternberg, R. J. (1981). Testing and cognitive psychology. The American Psychologist, 36, 1181-1189.

Sternberg, R. J. (1983). Components of Human Intelligence. Cognition, 15, 1-48.

Sternberg, R. J. (1985). (Ed.) Human abilities: An information-processing approach. New York: W. H. Freeman.

Sternberg, R. J. (1996). Cognitive Psychology. Fort Worth, Texas: Harcourt Brace College Publishers.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. American Psychologist, 38, 878-893.

Sternberg, S. (1966). High-speed scanning in human memory. Science, 153, 652-654.

Thorndike, R. L. (1973). Reading comprehension in fifteen countries. New York: Wiley.

Thurstone, L. L. (1938). Primary mental abilities. Psychometric Monographs, No. 1, 9-121.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. Neuropsychology, 11, 138-146.

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

Underwood, N.R., & Zola, D. (1986). The span of letter recognition of good and poor readers. Reading Research Quarterly, 21, 6-19.

Voss, J. F., Fincher-Kiefer, R. H., Greene, T. R., & Post, T. A. (1985). Individual differences in performance: The contrast approach to knowledge. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 3, pp. 297-334). Hillsdale, NJ: Erlbaum.

Woloshyn, V. E., Pressley, M., & Schneider, W. (1992). Elaborative-interrogation and prior-knowledge effects on learning facts. Journal of Educational Psychology, 84(1), 115-124.

Woloshyn, V.E., Willoughby, T., Wood, E., & Pressley, M. (1990). Elaborative interrogation influences adult learning of factual paragraphs. Journal of Educational Psychology, 82(3), 513-524.

Yekovich, F. R., Walker, C. H., Ogle, L. T., & Thompson, M. A. (1990). The influence of domain knowledge on inferencing in low-aptitude individuals. Psychological Learning and Motivation, 25, 175-191.

Appendix (continued)

	TRUE	FALSE
High	Like CHAIRS, GATHIS can be sat on. Like CHAIRS, SHETS can be sat on. Like CANDLES, MUNTŞ can't be sat on.	Like CANDLES, GATHIS can't be sat on. Like CANDLES, SHETS can't be sat on. Like CHAIRS, MUNTŞ can be sat on.
	<u>Knowledge Access Statements:</u>	
Low	A COUCH is heavier than a LAMP.	A LAMP is heavier than a COUCH.
High	A LAMP is used for light, whereas a CHAIR isn't. A CANDLE is used for light, whereas a COUCH isn't.	A CHAIR is used for light, whereas a LAMP isn't. COUCH is used for light, whereas a CANDLE isn't.

(1)

A NORT resembles a JET but is faster and weighs more.

A BERL. resembles a CAR but is slower and weighs more.

A SAMP resembles a BERL. but is slower and weighs more.

Features / Relations: speed NORT > JET > CAR > BERL > SAMP

weight NORT > JET > CAR ; SAMP > BERL > CAR

Text-Based Statements:

TRUE

Memory A NORT is faster than a JET.

FALSE

A JET is faster than a NORT.

Appendix (continued)

	TRUE	FALSE
	A BERL is slower than a CAR.	A CAR is slower than a BERL.
	A SAMP is slower than a BERL.	A BERL is slower than a SAMP.
	A NORT weighs more than a JET.	A JET weighs more than a NORT.
	A BERL weighs more than a CAR.	A CAR weighs more than a BERL.
	A SAMP weighs more than a BERL.	A BERL weighs more than a SAMP.
Inferencing	A SAMP is slower than a CAR.	A CAR is slower than a SAMP.
	A SAMP weighs more than a CAR.	A CAR weighs more than a SAMP.
<u>Knowledge Integration Statements:</u>		
Low	A NORT weighs more than a CAR.	A CAR weighs more than a NORT.
	A JET is faster than a BERL.	A BERL is faster than a JET.
Medium	A NORT is faster than a MOTORCYCLE.	A MOTORCYCLE is faster than a NORT.
	A ROCKET is faster than a BERL.	A BERL is faster than a ROCKET.
	A ROCKET is faster than a SAMP.	A SAMP is faster than a ROCKET.
High	Like ROCKETS, NORTS travel in the air.	Like MOTORCYCLES, NORTS travel across the land.
	Like MOTORCYCLES, SAMPs travel across the land.	Like ROCKETS, SAMPs travel in the air.
	Like MOTORCYCLES, BERLS travel across the land.	Like ROCKETS, BERLS travel in the air.

Appendix (continued)

TRUE

FALSE

Knowledge Access Statements:

Low	A JET is faster than a CAR.	A CAR is faster than a JET.
	A JET weighs more than a CAR.	A CAR weighs more than a JET.
High	A JET has a pilot, whereas a MOTORCYCLE doesn't.	A JET has a driver, whereas a MOTORCYCLE doesn't.
	A CAR has a driver, whereas a ROCKET doesn't.	A CAR has a pilot, whereas a ROCKET doesn't.

(2)

A ROSP resembles a DERP but is taller and lives longer.
 A DERP resembles a PINE but is taller and lives longer.
 A BUFT resembles a PETUNIA but is shorter and lives longer.

Features / Relations:

height	ROSP > DERP > PINE > PETUNIA > BUFT
lifespan	ROSP > DERP > PINE > PETUNIA; BUFT > PETUNIA

Text-Based Statements:

TRUE

FALSE

Memory	A ROSP is taller than a DERP.	A DERP is taller than a ROSP.
	A DERP is taller than a PINE.	A PINE is taller than a DERP.
	A BUFT is shorter than a PETUNIA.	A PETUNIA is shorter than a BUFT.

Appendix (continued)

	TRUE	FALSE
	A ROSP lives longer than a DERP.	A DERP lives longer than a ROSP.
	A DERP lives longer than a PINE.	A PINE lives longer than a DERP.
	A BUFT lives longer than a PETUNIA.	A PETUNIA lives longer than a BUFT.
Inferencing	A ROSP is taller than a PINE.	A PINE is taller than a ROSP.
	A ROSP lives longer than a PINE.	A PINE lives longer than a ROSP.
<u>Knowledge Integration Statements:</u>		
Low	A DERP lives longer than a PETUNIA.	A PETUNIA lives longer than a DERP.
	A PINE is taller than a BUFT.	A BUFT is taller than a PINE.
Medium	An EVERGREEN is taller than a BUFT.	A BUFT is taller than an EVERGREEN.
	A ROSP is taller than a TULIP.	A TULIP is taller than a ROSP.
	A DERP is taller than a TULIP.	A TULIP is taller than a DERP.
High	Like EVERGREENS, ROSPS have needles.	Like TULIPS, ROSPS have leaves.
	Like EVERGREENS, DERPS have needles.	LIKE TULIPS, DERPS have leaves.
	Like TULIPS, BUFTS have leaves.	Like EVERGREENS, BUFTS have needles.

Appendix (continued)

TRUE

FALSE

Knowledge Access Statements:

- | | | |
|------|--|---|
| Low | A PINE is taller than a PETUNIA.
A PINE lives longer than a PETUNIA. | A PETUNIA is taller than a PINE.
A PETUNIA lives longer than a PINE. |
| High | Sometimes PINES are decorated with ornaments, whereas TULIPS aren't.
Sometimes EVERGREENS are decorated with ornaments, whereas PETUNIAS aren't. (false)
Sometimes TULIPS are decorated with ornaments, whereas PINES aren't.
Sometimes PETUNIAS are decorated with ornaments, whereas EVERGREENS aren't. (false) | |
-

(3)

- A MIRT resembles an OSTRICH but is larger and has a longer neck.
- A COFT resembles a ROBIN but is smaller and has a longer neck.
- A FILP resembles a COFT but is smaller, has a longer neck, and nests on land.

Features / Relations:

size	MIRT > OSTRICH > ROBIN > COFT > FILP
neck length	MIRT > OSTRICH > ROBIN ; FILP > COFT > ROBIN
nests on land	FILP, OSTRICH
doesn't nest on land	COFT, ROBIN

Appendix (continued)

TRUE

FALSE

Text-Based Statements:

Memory A MIRT is larger than an OSTRICH.

An OSTRICH is larger than a MIRT.

A COFT is smaller than a ROBIN.

A ROBIN is smaller than a COFT.

A FILP is smaller than a COFT.

A COFT is smaller than a FILP.

A MIRT has a longer neck than an OSTRICH.

A MIRT has a longer neck than an OSTRICH.

A COFT has a longer neck than a ROBIN.

A ROBIN has a longer neck than a COFT.

A FILP has a longer neck than a COFT.

A COFT has a longer neck than a FILP.

A FILP nests on land.

A FILP doesn't nest on land.

Inferencing A FILP is smaller than a ROBIN.

A ROBIN is smaller than a FILP.

A FILP has a longer neck than a ROBIN.

A ROBIN has a longer neck than a FILP.

A COFT doesn't nest on land.

A COFT nests on land.

Knowledge Integration Statements:

Low A MIRT has a longer neck than a ROBIN.

A ROBIN has a longer neck than a MIRT.

An OSTRICH is larger than a COFT.

A COFT is larger than an OSTRICH.

Medium A MIRT is larger than a BLUEJAY.

A BLUEJAY is larger than a MIRT.

A PENGUIN is larger than a COFT.

A COFT is larger than a PENGUIN.

Appendix (continued)

TRUE

A PENGUIN is larger than a FILP.

Like PENGUINS, MIRTS can't fly.

Like BLUEJAYS, COFTS can fly.

Like BLUEJAYS, FILPS can fly.

Knowledge Access:

Low An OSTRICH is larger than a ROBIN.

An OSTRICH has a longer neck than a ROBIN.

A ROBIN nests in trees, whereas an OSTRICH nests on land.

High A ROBIN lives in Canada, whereas a PENGUIN typically doesn't.

A BLUEJAY lives in Canada, whereas an OSTRICH typically doesn't. (false)

A PENGUIN lives in Canada, whereas a ROBIN typically doesn't.

An OSTRICH lives in CANADA, whereas a BLUEJAY typically doesn't. (false)

FALSE

A FILP is larger than a PENGUIN.

Like BLUEJAYS, MIRTS can fly.

Like PENGUINS, COFTS can't fly.

Like PENGUINS, FILPS can't fly.

A ROBIN is larger than an OSTRICH.

A ROBIN has a longer neck than an OSTRICH.

An OSTRICH nests in trees, whereas a ROBIN nests on land.

(4)

A TOIP resembles a MARB but is more colourful, larger and lives in a colony.

A MARB resembles a BUTTERFLY but is more colourful and larger.

A JERP resembles an ANT but is less colourful and larger.

Appendix (continued)

Features / Relations: colorful TOLP > MARB > BUTTERFLY > ANT > JERP
 size TOLP > MARB > BUTTERFLY > ANT ; JERP > ANT
 lives in colonies: TOLP, ANT
 doesn't live in colonies MARB, BUTTERFLY

Text-Based Statements:

	TRUE	FALSE
Memory	A TOLP is more colourful than a MARB. A MARB is more colourful than a BUTTERFLY. A JERP is less colourful than an ANT. A TOLP is larger than a MARB. A MARB is larger than a BUTTERFLY. A JERP is larger than an ANT. A TOLP lives in a colony.	A MARB is more colourful than a TOLP. A BUTTERFLY is more colourful than a MARB. An ANT is less colourful than a JERP. A MARB is larger than a TOLP. A BUTTERFLY is larger than a MARB. An ANT is larger than a JERP. A TOLP doesn't live in a colony.
Inferencing	A TOLP is more colourful than a BUTTERFLY. A TOLP is larger than a BUTTERFLY. A MARB doesn't live in a colony.	A BUTTERFLY is more colourful than a TOLP. A BUTTERFLY is larger than a TOLP. A MARB lives in a colony.

Appendix (continued)

TRUE

FALSE

Knowledge Integration Statements:

Low	A MARB is larger than an ANT.	An ANT is larger than a MARB.
	A BUTTERFLY is more colourful than a JERP.	A JERP is more colourful than a BUTTERFLY.
Medium	A TOLP is more colourful than a COCKROACH.	A COCKROACH is more colourful than a TOLP.
	A MARB is more colourful than a COCKROACH.	A COCKROACH is more colourful than a MARB.
	A HONEYBEE is more colourful than a JERP.	A JERP is more colourful than a HONEYBEE.
High	Like HONEYBEES, TOLPS fly in the air.	Like COCKROACHES, TOLPS travel across the land.
	Like HONEYBEES, MARBS fly in the air.	Like COCKROACHES, MARBS travel across the land.
	Like COCKROACHES, JERPS travel across the land.	Like HONEYBEES, JERPS fly in the air.

Knowledge Access:

Low	A BUTTERFLY is more colourful than an ANT.	An ANT is more colourful than a BUTTERFLY.
	A BUTTERFLY is larger than an ANT.	An ANT is larger than a BUTTERFLY.
	ANTS live in a colony, whereas BUTTERFLIES don't live in a colony.	
	BUTTERFLIES live in a colony, whereas ANTS don't live in a colony. (false)	
High	ANTS have queens, whereas COCKROACHES don't	COCKROACHES have queens, whereas ANTS don't.
	HONEYBEES have queens, whereas BUTTERFLIES don't.	BUTTERFLIES have queens, whereas HONEYBEES don't.

Appendix (continued)

(5)

A WEMP resembles a GIRAFFE but is taller and eats more.

A long-legged TILN resembles a MONKEY but is shorter and eats more.

A LORK resembles a TILN but is shorter, eats more, and lives on land.

Features / Relations:	height	WEMP > GIRAFFE > MONKEY > TILN > LORK
	amount eaten	WEMP > GIRAFFE > MONKEY ; LORK > TILN > MONKEY
	long-legged	TILN, GIRAFFE
	short-legged	MONKEY
	lives on land	LORK, GIRAFFE
	doesn't live on land	TILN, MONKEY

Text-Based Statements:

	TRUE	FALSE
Memory	A TILN is long-legged.	A TILN isn't long-legged.
	A WEMP is taller than a GIRAFFE.	A GIRAFFE is taller than a WEMP.
	A TILN is shorter than a MONKEY.	A MONKEY is shorter than a TILN.
	A LORK is shorter than a TILN.	A TILN is shorter than a LORK.
	A WEMP eats more than a GIRAFFE.	A GIRAFFE eats more than a WEMP.

Appendix (continued)

	TRUE	FALSE
	A TILN eats more than a MONKEY.	A MONKEY eats more than a TILN.
	A LORK eats more than a TILN.	A TILN eats more than a LORK.
	A LORK lives on land.	A LORK doesn't live on land.
Inferencing	A LORK is shorter than a MONKEY.	A MONKEY is shorter than a LORK.
	A LORK eats more than a MONKEY.	A MONKEY eats more than a LORK.
	A TILN doesn't live on land.	A TILN lives on land.
	A LORK is long-legged.	A LORK isn't long-legged.

Knowledge Integration Statements:

Low	A WEMP eats more than a MONKEY.	A MONKEY eats more than a WEMP.
	A GIRAFFE is taller than a TILN.	A TILN is taller than a GIRAFFE.
Medium	A WEMP is taller than a LEOPARD.	A LEOPARD is taller than a WEMP.
	A BUFFALO is taller than a TILN.	A TILN is taller than a BUFFALO.
	A BUFFALO is taller than a LORK.	A LORK is taller than a BUFFALO.
High	Like BUFFALOS, TILNS are brown.	Like LEOPARDS, TILNS are spotted.
	Like BUFFALOS, LORKS are brown.	Like LEOPARDS, LORKS are spotted.
	Like LEOPARDS, WEMPS are spotted.	Like BUFFALOS, WEMPS are brown.

Appendix (continued)

TRUE

FALSE

Knowledge Access:

- Low A GIRAFFE is taller than a MONKEY. A MONKEY is taller than a GIRAFFE.
A GIRAFFE eats more than a MONKEY. A MONKEY eats more than a GIRAFFE.
A GIRAFFE lives on land, whereas a MONKEY lives in trees.
A GIRAFFE is long-legged, whereas a MONKEY is short-legged. (false)
A MONKEY lives on land, whereas a GIRAFFE lives in trees.
A MONKEY is long-legged, whereas a GIRAFFE is short-legged. (false)
High Typically, a GIRAFFE weighs a lot, whereas a LEOPARD doesn't.
Typically, a BUFFALO weighs a lot, whereas a MONKEY doesn't. (false)
Typically a LEOPARD weighs a lot, whereas a GIRAFFE doesn't.
Typically a MONKEY weighs a lot, whereas a BUFFALO doesn't. (false)

-
- (6) A GORT resembles a TESK but is larger, sweeter, and grows on a vine.
A round TESK resembles a WATERMELON but is larger and sweeter.
A PESH resembles a LEMON but is smaller and sweeter.

Appendix (continued)

<u>Features / Relations:</u>	size	GORT > TESK > WATERMELON > LEMON > PESH
	sweetness	GORT > TESK > WATERMELON > LEMON ; PESH > LEMON
	round	GORT, TESK
	not round	WATERMELON, LEMON
	grows on vine	WATERMELON
	doesn't grow on vine	TESK, LEMON

Text-Based Statements:

	TRUE	FALSE
Memory	A TESK is round.	A TESK isn't round.
	A GORT is larger than a TESK.	A TESK is larger than a GORT.
	A TESK is larger than a WATERMELON.	A WATERMELON is larger than a TESK.
	A PESH is smaller than a LEMON.	A LEMON is smaller than a PESH.
	A GORT is sweeter than a TESK.	A TESK is sweeter than a GORT.
	A TESK is sweeter than a WATERMELON.	A WATERMELON is sweeter than a TESK.
	A PESH is sweeter than a LEMON.	A LEMON is sweeter than a PESH.
	A GORT grows on a vine.	A GORT doesn't grow on a vine.
Inferencing	A GORT is larger than a WATERMELON.	A WATERMELON is larger than a GORT.

Appendix (continued)

	TRUE	FALSE
	A GORT is sweeter than a WATERMELON.	A WATERMELON is sweeter than a GORT.
	A GORT is round.	A GORT isn't round.
	A TESK doesn't grow on a vine.	A TESK grows on a vine.
	<u>Knowledge Integration Statements:</u>	
Low	A TESK is sweeter than a LEMON.	A LEMON is sweeter than a TESK.
	A WATERMELON is larger than a PESH.	A PESH is larger than a WATERMELON.
Medium	A GORT is larger than a GRAPEFRUIT.	A GRAPEFRUIT is larger than a GORT.
	A TESK is larger than a GRAPEFRUIT.	A GRAPEFRUIT is larger than a TESK.
	A CANTALOUPE is larger than a PESH.	A PESH is larger than a CANTALOUPE.
High	Like CANTALOUPE, GORTS aren't acidic.	Like GRAPEFRUITS, GORTS are acidic.
	Like CANTALOUPE, TESKS aren't acidic.	Like GRAPEFRUITES, TESKS are acidic.
	Like GRAPEFRUITS, PESHES are acidic.	Like CANTALOUPE, PESHES aren't acidic.
	<u>Knowledge Access:</u>	
Low	A WATERMELON is larger than a LEMON.	A LEMON is larger than a WATERMELON.
	A WATERMELON is sweeter than a LEMON.	A LEMON is sweeter than a WATERMELON.
	WATERMELONS grow on vines, whereas LEMONS grow on trees.	

Appendix (continued)

Typically, WATERMELONS and LEMONS aren't round. (false)

LEMONS grow on vines, whereas WATERMELONS grow on trees.

Typically, WATERMELONS and LEMONS are round. (false)

High Ripe WATERMELONS are green outside, whereas ripe GRAPEFRUITS aren't.

Ripe CANTALoupES are tan-coloured outside, whereas ripe LEMONS aren't. (false)

Ripe GRAPEFRUITS are green outside, whereas ripe WATERMELONS aren't.

Ripe LEMONS are tan-coloured outside, whereas ripe CANTALoupES aren't. (false)

Appendix B

Instructions for the New Component Processes Task.

We are interested in learning how people use what they already know to help them learn new information. When you press the spacebar to indicate that you are ready to proceed, the first sentence of a mini-paragraph will be presented on the computer screen. One of the terms in the sentence will be a real item which should be familiar to you. The other term, denoted by a nonsense syllable, will represent an imaginary item which is unfamiliar to you. Once you have learned this sentence press the spacebar for the next sentence of the mini-paragraph. Once you have learned all three sentences of the mini-paragraph you will be presented with some test statements.

IT IS IMPORTANT THAT YOU RESPOND QUICKLY AND ACCURATELY TO EACH TEST STATEMENT, BUT DON'T RESPOND SO QUICKLY THAT YOU MAKE A LOT OF ERRORS.

Note, that the information in a test statement can be true even though it was never explicitly stated in the paragraph. Your prior knowledge about the real items in the paragraph should enable you to deduce some relations that were not actually presented. Hence, some test statements should be considered true because the information in that statement was actually presented in the paragraph. Other test statements should be considered true because the information described in them could be deduced from your existing knowledge about real things in the world. When you have completed answering the test statements, the first sentence of the next mini-paragraph will appear on the computer screen.

DO YOU HAVE ANY QUESTIONS ???