

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Investigating the
relationship between two
approaches to verbal
information processing
in working memory :**

**AN EXAMINATION OF THE CONSTRUCT
OF WORKING MEMORY COUPLED WITH
AN INVESTIGATION OF META-
WORKING MEMORY**

Submitted in partial fulfilment for the degree of Doctor of Philosophy in Psychology
at Massey University by...

***Llewelyn A. Richards-Ward
Department of Psychology
Massey University
Palmerston North
New Zealand***

- 1996 -

ABSTRACT

Working memory is a process whereby persons can preserve information for a short time while concurrently engaging in other cognitive operations. The literature describes two approaches to working memory. The first approach (Baddeley, 1986) can be described as a complete model of working memory. However the second approach is not as clearly a distinct model, although its history, literature, application, simulation and operational definitions can arguably allow one to describe it as a separate model or strand of working memory for the present purposes. Rather, what will be termed the "quantitative/process model" deals only with verbal information and is far less complete than Baddeley's model in other domains. A central issue is thus how these two models relate with respect to how they handle verbal information. Baddeley (1986) delineated working memory as a set of interconnected components consisting of a Central Executive, a Phonological Loop, and a Visuo-Spatial Sketch-Pad. In this dissertation, this is termed the qualitative/structural model of working memory. Daneman and Carpenter (1980, 1983) delineated working memory as a process involving both a traditional span component and a concurrent operation. This approach, which will be referred to as the quantitative/process model of working memory, has been presumed to involve the Central Executive of the qualitative/structural model of working memory. This presumed relationship is scrutinised in the present dissertation in the context of an alternate hypothesis that the quantitative/process model involves more of the phonological loop than has been presumed. Thus, the first issue this dissertation addressed was how these two models or approaches to working memory account for verbal information. The second facet of the present investigation was to examine whether persons were able to report on their meta-memory for working memory.

Seven linked experiments are reported in the present dissertation. Participants for all seven experiments were predominantly students at local tertiary institutions and ranged in age from 16 to 48 years. The experimental conditions were presented as a two-factor within-subjects design in Experiments 1 to 6. The first general factor was word-type varying either across word-length (Experiments 2, 4, and 6) or across phonological similarity (Experiments 1, 3, and 5). The second factor was whether articulatory suppression was used or not (Experiments 1 to 6). In Experiments 1 and 2, stimuli were presented as a complex-span task (sentence plus word), where in Experiments 3 to 6, stimuli were presented as a simple-span task (word only). Experiment 5 also had a between-subjects factor determined by whether words were sampled from a 10 item pool or from a pool without replacement. Experiment 6 had a between-subjects factor determined by the presentation pace of the stimuli (at 1 per second or self-paced). Finally, Experiment 7 directly compared complex-span and simple-span presentations against a second factor of word-type varying across both phonological similarity and word-length (control, phonologically similar, 3-syllable).

In all seven experiments, participants were measured on dependent variables of recall in the correct serial position and recall in any serial position of the words that were presented. From the difference between these two measures of content, an estimate of the loss of order information (order errors) was calculated. A measure of the time each participant spent viewing (for simple-span tasks) or verifying (for complex-span tasks) the stimuli was made to assess processing time. Finally, before each trial, participants made an estimate of how many items they expected to recall in any order (a measure of their online meta-memory). In Experiments 5 to 7, a measure of the time each participant took to articulate the pool of words they had

been asked to recall was taken to provide an estimate of their articulation rate.

The main research questions for this set of studies were as follows: (1) that the quantitative/process model of working memory also uses the Phonological Loop, not just the Central Executive, and hence both models of working memory use the same process to preserve visually presented verbal information; (2) that measurement of dimensions of order and processing time, in addition to the dimension of content or capacity, will contribute independent information to the description of working memory function; and (3) that persons are able to monitor and report on their working memory. Data from the present set of studies provide support for these three hypotheses¹. The present investigation showed that a concurrent operation does not preclude phonological similarity and word-length effects used to define the components of the qualitative/structural model of working memory. Concurrently, dimensions of content and order, but not processing time, were shown to be important in describing working memory. The conclusion from these results is that both models of working memory refer to the same construct and that preservation of verbal information can be better accounted for by a single process. Finally, in all instances persons were accurate in predicting their general working memory performance. The data also show that persons may be able to predict the effect of some parameter changes on their performance.

¹The raw data upon which this dissertation was based can be obtained in the first instance from the author at the following address: Llewelyn A. Richards-Ward

c/- Department of Psychology
Massey University
Private Bag 11222
Palmerston North
New Zealand

The results of the present research suggest that verbal information is handled similarly in both models or approaches and tends to falsify that verbal information is retained primarily in the Central Executive in one model and the Phonological Loop in the other. Second the results suggest that persons do have a degree of meta-working memory. These results are discussed in terms of their implications for how working memory and meta-working memory can be described. Finally, some future directions for research are outlined.

ACKNOWLEDGEMENTS

No undertaking such as studying for a doctoral degree can occur without support and help from many people. However, to list all of those people who have supported me would be impossible. So, firstly I thank my friends who were there when I needed to talk, work colleagues who regularly asked how things were going, casual acquaintances who took the time to listen or offer advice, and those people who took the time to reply to my letters or email notes.

I would also like to thank Massey University for the doctoral scholarship which enabled me to undertake this research and for the Graduate Research Award A93/G/058 which helped cover some of the costs involved. My gratitude is also extended to the Psychology Department who contributed additional monies toward my research costs and to Michael Donnelly who helped me travel the paths of bureaucracy relatively unhindered.

I also thank Dr Ross St George and Dr Julie Bunnell for being my supervisors. They not only provided me with quality academic supervision but were also available to encourage and guide me in negotiating this degree with minimum trauma and maximum challenge. I remain very deeply indebted to them both.

At a more personal level, there are four people whose inspiration and support have, I believe assisted this undertaking greatly. First, I thank Ross St George for his enthusiasm and ability to ask questions. Without questions, research would never begin. Second, I thank Alison St George whose courses in educational psychology inspired me to consider more deeply how people monitor their cognition and helped me develop my own strategies for learning greatly. Third, I thank John Podd for teaching me the necessity of taking small steps in research and of the value of being thorough. Fourth, I thank Julie Bunnell whose teaching first

inspired my interest in cognitive psychology and whose support has enabled me to develop that interest. To all four of you, please do not let the brevity of this acknowledgement of your contributions in any way detract from the appreciation I extend toward you.

Finally, I wish to thank Leigh without whose participation as an (in)volunteer and support as a partner I could never have completed this degree. *"A good wife who can find? She is far more precious than jewels. The heart of her husband trusts in her, and he will have no lack of gain"* (Proverbs 31:10-11). Diolch yn fawr rwyn dy cariad di Leigh bach!

TABLE OF CONTENTS

Abstract	i
Acknowledgements.....	v
Table of Contents	vii
List of Tables.....	xii
List of Figures.....	xvi

Introduction.....	1
--------------------------	----------

Chapter 1: A History of the Short-Term Store and of Working Memory.....	3
What is Memory?: Early Models of Memory	5
The Span Method of Measuring the Capacity of the Short-Term Store.....	6
The Free-Recall Method of Measuring the Capacity of the Short-Term Store	7
A Comparison of the Short-Term Store Using Either Span or Free-Recall Measurement Techniques	8
What is the Short-Term Store for?: The Baddeley and Hitch (1974) Investigation of Working Memory	10
Summary	12
Chapter 2: Qualitative/Structural Models of Working Memory	14
An Overview of Baddeley's Definition of Working Memory.....	15
The Qualitative/Structural Model of Working Memory	16
The Phonological Loop.....	17
The Visuo-Spatial Sketch-Pad	32
The Central Executive.....	33
Summary	35

Chapter 3: Quantitative/Process Models of Working Memory	37
A Pragmatic Analysis of the Complex-Span Tasks: Do They Perform?.....	38
Definition of Working Memory	40
Reliability of Complex-span tasks	44
Validity Evidence for the Reading Span task.....	45
Conclusions	51
A Theoretical Framework for the Quantitative/Process Model of Working Memory	51
Preservation of Content in Working Memory	52
Order Information in Working Memory	53
Temporal Factors in Working Memory	54
Conclusions	56
Summary	56
Chapter 4: Subjective On-line Meta-Memory Measures	58
The Theoretical Basis of Meta-Memory.....	60
The Operational Basis of Meta-Memory	64
Assumptions of Meta-Memory Research.....	67
Conclusions	68
Chapter 5: A Rationale for Investigating the two Models of Working Memory	69
Relationships Between the Qualitative/Structural and Quantitative/Process Models of Working Memory.....	73
The Relationship Between Meta-Memory and Working Memory	79
Toward an Integration of Theories and Measurement of Verbal Working Memory	80
Summary	81

Experimentation	83
------------------------------	-----------

Chapter 6: General Hypotheses and General Method	83
General Hypotheses: Three Defining Postulates	83
Hypotheses at a Theoretical Level	84
General Method	89
Participants	89
Apparatus.....	90
Stimuli	90
Design	93
Procedure	95

Chapter 7: Experiments 1 & 2: An Initial Examination of Quantitative/Process, Qualitative/Structural, and Subjective/Objective Factors in a Complex-Span Task	102
Experiment 1: An Initial Test of the Phonological Similarity Effect in a Complex-Span Task	104
Method	108
Results	109
Discussion	120
Experiment 2: The Word-length Effect in a Complex-span Task	127
Method	133
Results	134
Discussion	142
Summary of Experiments 1 and 2	147
Chapter 8: Experiments 3 & 4: Examination of Quantitative/Process, Qualitative/Structural, and Subjective/Objective factors in a Simple-Span Task	150
Experiment 3: The Phonological Similarity Effect in a Simple-Span Task	151
Method	153
Results	154
Discussion	160
Experiment 4: The Word-length Effect in a Simple-span Task	165
Method	167
Results	168
Discussion	173
Summary	176
Utilizing Dimensions of Order and Processing Time in Addition to Content Information	176
Meta-memory for Working Memory	177
Conclusions	178
Chapter 9: Replicating the Phonological Similarity and Word-Length Effects in a Simple-Span Task	179
Experiment 5: A Comparison of Two Stimulus Sampling Methods on the Phonological Similarity Effect	180
Method	194
Results	195
Discussion	210
Experiment 6: A Comparison of the Effect of Experimenter-Paced Versus Participant-Paced Presentation of Stimuli on Word-length Effects	214
Method	221
Results	222
Discussion	230

Summary of Experiments 5 and 6.....	233
Chapter 10: Experiment 7: The Integration of a Qualitative/Structural Model of Working Memory with a Quantitative/Process Model of Working Memory	236
Hypotheses.....	241
Method	243
Results	245
Discussion	255

Discussion and Conclusions	262
---	------------

Chapter 11: General Discussion.....	262
Pre-conditions for Measurement of Working Memory:	
Reliability, Sensitivity, and Bias	263
Reliability	263
Sensitivity.....	264
Bias	265
Postulate 1: The Commonality Between Qualitative/Structural and Quantitative/Process Definitions of Working Memory.....	265
Main Results	267
The Context of The Present Results in the Literature.....	271
Conclusion.....	273
Postulate 2: The Utility of Measures of Content, Order and Processing Time in Describing Working Memory.....	274
Main Results	274
Limits of the results.....	279
Conclusion.....	283
Postulate 3: On-line Meta-Memory for Working Memory Content.....	284
Main Results	284
Conclusion.....	288
Summary	289
The Unity of the Construct of Working Memory.....	283
Implications of the Present Results for Theories of Working Memory	284
Implications of the Present Results for Meta- Memory Research	293
Future Research	296
Conclusion.....	299

References	301
Appendices	318
Appendix 1: Experimental Stimuli Used in the Present Dissertation.	318
Appendix 2: A Signal Detection Analysis of the Sensitivity and Bias of Participants' Responding on the Sentence Verification Task	326
Appendix 3: The General Information Sheet for Experiments 1 to 7.....	328
Appendix 4: The Participant Consent Sheet for Experiments 1 to 7.....	329
Appendix 5. F-values and errors terms for Experiment 7.....	330

LIST OF TABLES

Table 1.	A comparison of the span and free-recall methods of measuring STS capacity	9
Table 2.	A selective review of working memory definitions used by researchers using a quantitative and process oriented correlational approach to aging and working memory	39
Table 3.	Reliability coefficients for various operationalisations of a storage plus processing models of working memory.	44
Table 4.	Reported correlations between the Daneman and Carpenter (1980) Reading Span task (or a variant) and other measures of working memory or of memory span alone.	45
Table 5.	A summary of the experimental conditions used in the present dissertation. Conditions vary across phonological similarity, word-length, articulatory suppression, tasktype, pool-type, and pacing	91
Table 6.	A summary of the two week test-retest reliability of the complex-span task for phonologically similar and dissimilar words with and without articulatory suppression in Experiment 1. Test-retest reliability was calculated using the Pearson product-moment coefficient	111
Table 7.	Mean (SD) recall of phonologically similar or phonologically dissimilar words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions in Experiment 1. Note that recall is out of three for either correct or muddled verification conditions and out of 6 for the total verification and pre-estimates columns (N = 21).	116
Table 8.	Mean (SD) order errors for phonologically similar or phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 1. Order errors are out of 3 for the correct and muddled conditions, and out of 6 overall (N = 21)	117
Table 9.	Mean (SD) verification times (ms) for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 1 (N = 21)	119

Table 10. Mean (SD) pre-estimates of recall in any serial position for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 1 (N = 21)	120
Table 11. A summary of the two week test-retest reliability of the complex-span task for 1 and 2-syllable words with and without articulatory suppression in Experiment 2. Test-retest reliabilities were moderate to high in all conditions and were calculated using the Pearson product-moment coefficient.....	135
Table 12. Mean (SD) recall of 1-syllable or 2-syllable words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions in Experiment 2. Note that recall is out of three for either correct or muddled verification conditions and out of 6 for the total verification and pre-estimates columns (N = 21).	137
Table 13. Mean (SD) order errors for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 2. Order errors are out of 3 for the correct and muddled conditions, and out of 6 overall (N = 21)	139
Table 14. Mean (SD) verification times (ms) for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 2 (N = 21)	140
Table 15. Mean (SD) pre-estimates of recall in any serial position for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 2 (N = 21)	141
Table 16. Mean (SD) recall of phonologically similar or phonologically dissimilar words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions in Experiment 3 (N = 20).	155
Table 17. Mean (SD) order errors for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 3 (N = 20)	157
Table 18. Mean (SD) viewing times (ms) for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 3 (N = 18).	159

Table 19. Mean (SD) pre-estimates of recall in any serial position for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 3 (N = 18)	159
Table 20. Mean (SD) recall of 1-syllable or 2-syllable words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions in Experiment 4 (N = 19).....	168
Table 21. Mean (SD) order errors for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 4 (N = 19).....	170
Table 22. Mean (SD) viewing times (ms) for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 4 (N = 17).....	171
Table 23. Mean (SD) pre-estimates of recall in any serial position for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 4 (N = 18)	172
Table 24. Mean (SD) recall of phonologically similar or phonologically dissimilar words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions and across pool-type in Experiment 5 (N = 34).	199
Table 25. Articulation rate (SD) for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions and across pool-type in Experiment 5 (N = 34).....	201
Table 26. The proportion of variability accounted for in the Fixed-pool condition using the natural logarithm of each of the content measures regressed against the natural logarithm of articulation rate with <i>F</i> -values, significance, and mean square errors of the residual terms evaluating the fit of the regression coefficient	202
Table 27. Mean (SD) order errors for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions and across pool-type in Experiment 5 (N = 34).....	204
Table 28 Mean (SD) viewing times (ms) for both phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions and across pool-type in Experiment 5 (N = 34).....	207

Table 29. Mean pre-estimates (SD) for both phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions and across pool-type in Experiment 5 (N = 33).....	208
Table 30. Mean (SD) recall of 1-syllable or 3-syllable words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions across experimenter and participant-paced presentations in Experiment 6 (N = 34).	223
Table 31. Articulation rate (SD) for 1-syllable or 3-syllable words for both the control and articulatory suppression conditions and across pacing conditions in Experiment 6 (N = 34)	226
Table 32. Mean (SD) order errors for 1-syllable and 3-syllable words for both the control and articulatory suppression conditions and across the pacing conditions in Experiment 6 (N = 34).	226
Table 33. Mean (SD) viewing time (ms) for both 1-syllable and 3-syllable words for both the control and articulatory suppression conditions and across pacing conditions in Experiment 6 (N = 34)	228
Table 34. Mean (SD) pre-estimates of recall in any serial position for both 1-syllable and 3-syllable words for both the control and articulatory suppression conditions and across pacing conditions in Experiment 6 (N = 33).	229
Table 35. A summary of the domains of previous experiments in the present research.	237
Table 36. Mean (SD) recall of control, phonologically similar, and 3-syllable words either in the correct serial position or in any serial position by task-type (simple or complex) for Experiment 7 (N = 32).	246
Table 37. Mean (SD) order errors by task-type (simple or complex) and stimulus-type (control, phonologically similar, 3-syllable; N = 32) for Experiment 7.....	250
Table 38. Mean (SD) viewing and verification time (ms) by task-type (simple or complex) and stimulus-type (control, phonologically similar, 3-syllable; N = 32) for Experiment 7.	252
Table 39. Mean (SD) pre-estimates by task-type (simple or complex), and stimulus-type (control, phonologically similar, 3-syllable; N = 32) for Experiment 7.....	254

LIST OF FIGURES

Figure 1. A model of working memory showing the Central Executive, the Phonological Loop, and the Visuo-Spatial Sketch-Pad.	16
Figure 2. Salthouse's (1992) model of the relationship between storage and operations in working memory.	43
Figure 3. A flowchart showing some aspects of the logic behind the development of the present studies in relation to Postulate 1.....	87
Figure 4. Recall in the correct serial position by phonological similarity, articulatory suppression, and verification conditions for Experiment 1.....	112
Figure 5. Recall in any serial position by phonological similarity, articulatory suppression, and verification conditions for Experiment 1.....	115
Figure 6. Recall in the correct serial position by word-length, articulatory suppression, and verification conditions for Experiment 2.	136
Figure 7. Recall in the correct serial position and recall in any serial position by phonological similarity and articulatory suppression conditions in Experiment 3.....	156
Figure 8. Recall in the correct serial position and recall in any serial position by word-length and articulatory suppression conditions in Experiment 4.	169
Figure 9. Recall in the correct serial position by phonological similarity, articulatory suppression, and pool-type conditions in Experiment 5.	196
Figure 10. Recall in any serial position by phonological similarity, articulatory suppression, and pool-type conditions in Experiment 5.	198
Figure 11. A comparison of the two measures of content in Experiment 5 with each other.....	200
Figure 12. Item transpositions for phonologically similar and phonologically dissimilar words in Experiment 5.....	206
Figure 13. Recall in the correct serial position by word length, articulatory suppression, and pacing conditions in Experiment 6.	225
Figure 14. Item transpositions for 1- and 3-syllable words in Experiment 6	227

Figure 15.	Recall for Experiment 7 showing the comparison between simple-span and complex-span tasks for each stimulus-type.	248
Figure 16.	The natural logarithm of viewing times of each stimulus plotted against recall in the correct serial position for each task-type and for each stimulus-type.	253
Figure 17.	A regression of mean recall in any serial position against mean pre-estimates for each condition in Experiments 1 to 7.	285
Figure 18.	The qualitative/structural and quantitative/process models of working memory	292

Introduction

Chapter 1 presents the theoretical background and history of the working memory construct and how it was perceived to be a more effective model than a short-term store (STS). As was demonstrated by Baddeley and Hitch (1974), either multiple modality-specific STS's were to be proposed, or a new model of a short-term and transient memory needed to be described. They described a new model of *working memory*, consisting of structures in which certain qualitative effects occurred. Hence, this model is referred to as the qualitative/structural model of working memory.

In Chapter 2, relevant research and theory is presented to outline the strengths and flaws of the qualitative/structural working memory model. Briefly, the qualitative/structural model provides a good account of many areas of research in the working memory domain. However, the main deficiency of the qualitative/structural model of working memory is its reliance on recall as the only measure of working memory's products.

In Chapter 3, the quantitative/process model of working memory will be described. Evolving from the development of the qualitative/structural model of working memory, two other areas of applied cognition are reviewed that also make use of the concept of working memory: the area of reading (e.g., Daneman & Carpenter, 1980) and the area of aging (e.g., Salthouse, Mitchell, Skovronek, & Babcock, 1989). In both of these areas of applied cognition, working memory was defined as consisting of an *operation plus storage*. This definition allowed researchers to derive operationalisations of working memory sufficiently robust to predict both reading comprehension and age-related decrements in working memory. Because these operation plus storage models primarily obtained a score or quantity that represented the capacity of

working memory, and because they refer to working memory as a process without making inferences about structure, this group of working memory models will be referred to as representing a quantitative/process model of working memory. While the major strength of the quantitative/process model of working memory has been its ability to relate measures of working memory to measures of other relevant cognitive constructs, it suffers from a lack of theoretical depth -- a depth that the qualitative/structural model appears to have.

Chapter 4 examines a third dimension of working memory, defined by the distinction between measuring working memory by overt performance and by self-report. This distinction has been captured in the past by referring to performance and meta-memory measures respectively. The major issue examined in Chapter 4 is whether persons might possess an on-line meta-memory for specific working memory components in addition to a more general meta-memory for recall. The aim of Chapter 4 is to present a theoretical and empirical basis for the concept of meta-memory so that it can then be more meaningfully compared to working memory performance.

At the end of Chapter 4, three dimensions or ways of measuring working memory will have been discussed. At this point, it will be apparent that very little overlap at both an operational and theoretical level exists between these dimensions. This lack of overlap begs two questions that this dissertation examines in relation to how verbal information is processed: *Are the two models of working memory in fact examining the same construct?* and *Can persons monitor their working memory?* Chapter 5 presents the rationale for examining the relationships between the two models of working memory, and between working memory and meta-memory.

CHAPTER 1: A HISTORY OF THE SHORT-TERM STORE AND OF WORKING MEMORY

Memory is a core concern in the science of psychology. Ebbinghaus (1885/1964), in perhaps the first monograph published in experimental psychology, examined memory. Subsequent psychological research has, up to the present day, continued to be concerned with memory. Even those researchers with a behavioural focus, researchers who largely rejected the examination of cognition, developed associationist theories about memory.

To understand the importance of memory, one must consider what memory is. The word memory is derived from the Latin term *memor*, a term loosely translated into English as meaning mindful. Memory is connected to the mind. One can speculate whether memory is what provides an awareness of the conscious present (Crowder, 1993), whether one can learn without memory, whether learning and memory are the same, whether memory is necessarily something we can be aware of, and many other semi-philosophical questions. In doing so, it becomes clear that the core features of memory are that it allows retention of an organism's perception of the environment for periods of time ranging from milliseconds to a lifetime, a retention that is effected solely by the organism's biological processes. In doing so, memory provides a bridge to cross the boundary from perception to cognition.

As psychology began to expand beyond behavioural theories in the 1960's, there was a revived concern in understanding what memory is. As researchers considered the nature of memory, a dichotomy began to emerge between a shorter-term and a longer-term memory. A Short-Term Memory (STM) was considered to be transient (limited to a few seconds), primarily acoustic in nature, and served the function of preserving one's perceptions for long enough periods of time to allow other cognitive processing to operate on that memory. In contrast, a Long-Term Memory (LTM) was considered to be relatively permanent, to be primarily semantic, and served

the function of preserving the products of one's cognitive processing.]

The focus of the present dissertation is on the transient preservation of information. Throughout this dissertation, two terms are used to refer to this transient memory and, as such, need definition. First, the present research uses the term Short-Term Store (STS) to describe a structure inferred to preserve perceptions of the environment for a brief period of time. The STS is typically studied in the context of a triplex model of memory in which the three major components are a very short-term iconic store, a STS, and LTM (e.g., Klatzky, 1980). When the term STS is used in the present dissertation, there is an implied focus on the capacity of this transient memory, that is, *how big the transient memory is*.

The second term used to refer to this transient preservation of information is working memory. The concept of working memory is similar to that of the STS in that working memory also refers to a transient memory. However, as will be developed further below, the term *working memory* reflects the idea that this transient memory is used for later cognitive processing. That is, working memory is also described in terms of its *relationship to other cognitive processes* beyond memory processes alone. Working memory also involves description of how the transient STS operates. That is, working memory is about *how the STS preserves information* for short periods of time and about the parameters that influence that preservation.

In conclusion, it must be noted that the distinction between the STS and working memory has varied from one researcher to another. Some (e.g., Swanson, 1994) use the terms to refer to separate entities, whereas others (e.g., Baddeley, & Hitch, 1974) use them almost synonymously. Thus, it seems virtually impossible to present a clear distinction between the STS and working memory based on existing literature. However, the present dissertation uses the term working memory when it is intended that the reader understand that information is being

preserved for some other cognitive processing and that the nature of this transient preservation involves more than simply how much information can be preserved. In this regard, the present dissertation implies that the STS is subsumed in the term working memory.

While both the nature and the functions of working memory have been examined previously, the use of the term working memory has not necessarily been based on a consistent definition of how this working memory operates. That is, the present dissertation is concerned with comparing two predominant models of working memory with each other. In comparing differing models with each other, the aim of the present research was to begin to describe, in a more integrated fashion, how working memory operates.

However, to understand working memory, it is first necessary to understand something of its predecessor the STS. To this end, the first section of Chapter 1 presents a review of the two main ways of measuring the STS: by using a serial-recall span task and by using a free-recall task. The major flaw with both of these methods was that span did not appear to be related to any other cognitive processes, begging the questions of what the STS was for and how it operated (Baddeley & Hitch, 1974).

The second section of Chapter 1 reviews a seminal study by Baddeley and Hitch (1974) who, in beginning to define what the STS was for, cemented the idea in the minds of psychologists that the STS performed 'work' for other cognitive processes and consisted of an operation plus a STS or span. Thus, the concept of working memory became part of the popular parlance of applied psychology.

WHAT IS MEMORY?: EARLY MODELS OF MEMORY


There has been a tradition of defining short-term storage in terms of the tasks used to measure the capacity of that Short-Term Store (STS). The following section examines the span and free-recall methods, respectively, of measuring STS

capacity. Although this review is not intended to be a complete history of the STS, it will be apparent that how the STS was described in large part depended on the measurement method used. It is perhaps not coincidental that working memory has also been described in part by the measurement method used. This point will be expanded upon in Chapters 2 and 3.

Finally, a comparison is made between span and free-recall methods of measuring the STS. Overall, the conclusion of the following section is that while both the span task and the free-recall task measure the capacity of the STS, they are unable to address how the STS preserves content information, or what the STS is used for in relation to other cognitive activities.

The Span Method of Measuring the Capacity of the Short-Term Store

The key features of the span method of assessing STS capacity are as follows: (1) a fixed number of items are presented sequentially; (2) those items are to be recalled in the correct order; and (3) stimulus parameters vary across presentation modality, the use of interfering tasks between stimulus presentations and during retention interval, and the addition of a suffix after presentation but before recall. The capacity of the STS, using the span method, is typically defined as the number of items that were presented in a trial when all of those items were perfectly recalled. For example, if a person was presented trials of 4, 5, and 6 items and recalled all of the items from trials with 4 and 5 items but not with 6, a strict capacity score would record their span as 5 items.

Using the span method of assessing STS capacity, researchers proposed that there are three primary features of the STS. These features were also shown in each instance to be true for the STS, but not for LTM. Thus, the STS was very much defined in relation to LTM. First,  the STS has been shown to be limited in capacity to about 7 items (Ebbinghaus, 1885/1964; Miller, 1956). Second, the STS has been shown to

be subject to rapid decay without rehearsal (Brown, 1958; Conrad, 1967; Peterson & Peterson, 1959). Finally, the STS has been inferred to be acoustic in nature (Atkinson & Shiffrin, 1968, 1971) as defined by the acoustic similarity effect (Conrad, 1964; Baddeley, 1968; Laughery, 1969; Murray, 1968; Watkins, Watkins, & Crowder, 1974), the modality effect (Bigham, 1894, cited in Murray, 1980; Corballis, 1966), and the suffix effect (Crowder, 1967; Crowder & Morton, 1969; Greene 1992). However, this account of the STS was not the only one. It was being developed at the same time as the method of free-recall was being used to define the STS.

The Free-Recall Method of Measuring the Capacity of the Short-Term Store

Like the span method of assessing the STS, the free-recall method also measured the capacity of the STS. However, in the free-recall method, the STS is believed to produce the recency effect of the serial position curve (Craik, 1970), with LTM producing the primacy effect of the serial position curve (Murdock, 1962). Consequently, capacity is calculated from the recency portion of the serial position curve. However, unlike the span method, the free recall method uses a stochastic method, rather than an absolute capacity measure, to quantify STS capacity (Waugh & Norman, 1965).

Initially, the recency effect was assumed to occur because of the more recent items being preserved in the STS with the others decaying (Peterson & Peterson, 1959). However, Waugh and Norman (1965) showed that recall of items in a list was a function of the number of interfering items. This result presented a major difficulty for the decay theory of the STS. Of major importance to the present discussion was the finding that when there were few interfering items, recall was quite high. From these results, it was proposed that recall from the STS is limited by the amount of *proactive interference* operating upon each serial position.

A second feature of the free-recall method was that items at the start of a list were more likely to be confused with

semantically similar items and that items at the end of the list were more likely confused with acoustically similar items (Kintsch & Buschke, 1969; Tell, 1972). From these results, it was proposed that the STS was represented *acoustically*, and LTM was represented *semantically*.

There are, however, some data that suggest that the recency effect may not necessarily be produced by a transient (over a few seconds) STS, hence inferences made about the STS from the recency effect may be incomplete or incorrect. That the recency effect is not solely produced by the last few seconds of information a person receives can be shown by three effects: the effect of *concurrent distraction* (Baddeley & Hitch, 1977; Murdock, 1965); recall of *multi-category lists* (Watkins & Peynircioglu, 1983); and the *continuous-distractor* paradigm (Bjork & Whitten, 1974; Tzeng, 1973). These three effects demonstrate that recency effects occur over period much longer than the few seconds a STS supposedly maintains information for. Hence, if recency effects can occur over quite long periods of time, it cannot be inferred that they are produced solely by a STS whose duration is assumed to be only a few seconds.

A Comparison of the Short-Term Store Using Either Span or Free-Recall Measurement Techniques

At this point, two methods of measuring STS capacity have been briefly described along with some research relevant to those methods. From the evidence thus far, it would appear that the span task provides the most unambiguous account of the capacity of the STS. However, neither approach is without its faults.

Table 1 shows that the main points of difference between the span and free-recall methods of measuring the STS are that capacities are different between the two methods, and that recency is not limited to simply seconds of time, but can occur days after presentation. These differences suggest that the span and free-recall methods may in fact measure separate phenomena (Klatzky, 1980).

Table 1. A comparison of the span and free-recall methods of measuring STS capacity.

Span measures of the STS	Free-recall measures of the STS -- the recency effect
An acoustic similarity effect is present (e.g. Conrad & Hull, 1964; Watkins, et al., 1974).	No acoustic similarity effect occurs in the recency portion (Watkins, et al., 1974).
Word-frequency effects occur but are dependent on list position (e.g., Watkins, 1977).	Variables of list-length, word-frequency, and presentation rate affect pre-recency items. No similar dissociations exist for recency items (Greene, 1992).
STS span is empirically based.	STS span is formula based (e.g., Waugh & Norman, 1965).
Order information is confounded with content information.	Order information is separate from content information.
Categories or 'chunks' do not affect capacity (Miller, 1956).	Multiple categories presented together each show a recency effect (Watkins & Peynircioglu, 1983).
Span estimates typically range from 5-9 items.	Span estimates typically range from 2-4 items.
A modality effect occurs for the last items in the span (Corballis, 1966; Greene, 1992).	Recency effects are not confined to a short temporal duration, but can occur over days or weeks (Bjork & Whitten, 1974; Tzeng, 1973).
A suffix effect occurs with an auditorily presented suffix at the end of a span (Crowder, 1967).	The recency portion is unaffected by a concurrent memory span task (e.g., Baddeley & Hitch, 1974; Murdock, 1965).

Thus, at this point in history, the question of how large the capacity of the STS was had been answered by examining scores on either the span or free-recall procedures. Although some useful research had been done, the dual-store approach was clearly inadequate at many levels. As it became clear that neither the free-recall nor span methods of measuring capacity were completely adequate explanations of the STS, they were gradually replaced by the levels-of-processing approach (Craik & Tulving, 1975) and later by connectionist approaches (e.g., Schneider & Detweiler, 1988). The move beyond examination of STS capacity per se also represented, in some respects, a change in research question. At the start of this section, it was noted that the STS models were an answer to the question, "what is the capacity of a shorter term memory?" In contrast, moves to levels of processing and connectionist approaches were steps towards answering the question addressed in the present dissertation: the question of how short term information is maintained. Asking what

memory is assumes that memory is an entity. Asking how memory works allows a presupposition that memory may be a part of an overall cognitive process and may have dimensions other than capacity alone. For example, when one reads research upon which the model of the STS was based, it is also apparent that the STS is concerned with both content and order information, despite the focus on measuring capacity alone (e.g., Conrad, 1965; Ebbinghaus, 1885/1964; Estes, 1972; Healy, 1974, 1982; Lewandowsky & Murdock, 1989; Shiffrin & Cook, 1978).

However, in 1974, Alan Baddeley and Graham Hitch asked an intermediary question -- what is this STS for? This question was intermediary in that it did not entirely abandon seeking the answer to the question of what memory is through an analysis of the capacity of the STS. Asking what memory is for also introduced the idea that (1) memory may be contextualised within a larger cognitive arena and that (2) it may not necessarily be an entity, but may be a process. The answers to the question of what memory is for have revolutionized memory research by contextualising the STS as having a function in the operation of other cognitive processes and by directing research away from a narrow focus on the size of the STS.

WHAT IS THE SHORT-TERM STORE FOR?: THE BADDELEY AND HITCH (1974) INVESTIGATION OF WORKING MEMORY

Baddeley and Hitch (1974) conducted a seminal group of studies designed to examine whether STS resources are critical in storing the intermediate products of reasoning, comprehension, and learning. That is, they were asking what the role of a transient memory was in those tasks -- what memory was for.

The basic premise of the research by Baddeley and Hitch (1974) was that if the STS was essential for all other cognitive operations, then using that store will produce performance declines in other tasks. Conversely, as they noted, the evidence favoured the reverse conclusion that

other cognitive operations actually occurred in parallel with the operation of the STS. In a related vein, Baddeley and Hitch also asked if the STS played the role ascribed it by Atkinson and Shiffrin (1971) as not only a store, but a controller of other cognitive processes. That is, is the STS the same as a *working memory*?

Baddeley and Hitch (1974) conducted a series of studies in which participants were required to learn lists of varying lengths (pre-loading the STS) for later recall immediately prior to concurrently engaging in either a reasoning, comprehension, or free-recall task. With a pre-load list length of three items, Baddeley and Hitch found that performance on the primary and concurrent task was relatively unimpaired. For example, in their eighth experiment examining learning, Baddeley and Hitch required participants to recall (immediately or delayed) a sequence of 16 unrelated words in the correct serial order. This procedure typically shows a serial position effect wherein the most recently presented items are recalled in ~~greater~~ frequency than other items (e.g., Kausler, 1991). Prior to engaging in the free-recall task, participants were presented with 0, 3, or 6 digits to recall -- a *dual-task procedure*. Because the STS is in part defined by the serial position recency effect, presenting three digits was expected to have replaced about half of the available STS capacity for the words and to have decreased the recency effect for the most recently presented words in the free-recall task. Similarly, 6 digits were expected to have almost eliminated the recency effect after recall of the digits. What did happen was that a pre-load of 3 digits had almost no effect on the participants' ability to recall words, either immediately after the digits, or after a delay period. Six digits did disrupt the recency effect a little, but the disruption was not great. These data were typical of the degree of disruption that pre-loading the STS had on concurrent reasoning and comprehension tasks also, with one important exception. In both the moderately difficult reasoning and comprehension tasks, the effect of a pre-load

of 6 digits did produce a noticeable decrement in the other concurrent task. The data were consistent with the interpretation that the capacity of a working memory only in part consists of what has been termed the STS and that filling the STS does not preclude other concurrent cognitive operations, even intuitively related ones.

In proposing their initial formulation of working memory, Baddeley and Hitch (1974) suggested that working memory

...appears to have something in common with the mechanism responsible for the digit span, being susceptible to disruption by a concurrent digit span task, and like the digit span showing signs of being based at least in part upon phonemic coding. It should be noted ... that the degree of disruption observed...was far from massive.
(p.75)

To explain the less than massive disruption of a concurrent operation by the STS task, Baddeley and Hitch concluded that the management of a cognitive operation did not use storage capacity alone, rather storage and operation management were separate but related. Also, they suggested that there may be more than one type of storage in working memory, and the idea of a controller plus slave systems, while not described, was obviously the next logical inference. This idea concurs with the Atkinson-Shiffrin (1971) distinction between storage and control processes. The key point from the above discussion of the Baddeley and Hitch study is that working memory involves more than storage alone. The idea of a working memory space in which divided storage and control operations both occurred was (re)born.

SUMMARY

The present chapter began with a review of how the STS has been described in the past, which has been in terms of capacity. This review showed that there were two, perhaps related but not identical, ways of measuring STS capacity. Neither the free-recall nor span methods of measuring the STS

provided complete descriptions of the STS. At this point, it is argued that the research question changed. Baddeley and Hitch (1974) stated that:

Despite...the vast amount of research on the characteristics of the STS, there is still little general agreement. If our subsequent work were to depend on a generally acceptable definition of STS as a prerequisite for further research, such research would never begin.
(p.49)

In response to this conclusion, they asked a new question by asking what a transient memory was for. Baddeley and Hitch found that the answer to their question suggested that short-term maintenance of information involved both storage and control processes.

This chapter has shown that working memory is a construct used to define short-term preservation and management of information, in order that other cognitive processes might use that information. Baddeley and Hitch (1974) advanced the idea that working memory involved an operation plus what had hitherto been termed a STS. The two succeeding chapters examine how this operation plus span distinction proposed by Baddeley and Hitch (1974) was separately developed into a qualitative/structural model and a quantitative/process model of working memory by different groups of researchers.

CHAPTER 2: QUALITATIVE/STRUCTURAL MODELS OF WORKING MEMORY

Before discussing the qualitative/structural model, it is important to outline why this author has chosen to name it as such. Baddeley and others most often refer to this model as the tripartite model of working memory (Baddeley, 1986, 1992a, 1992b). However, this name implies three components. This dissertation explicitly rejects that working memory need necessarily be described in terms of components and, even if working memory was described using components, that there are only three of 'them'. In labelling this model of working memory as a qualitative/structural model two ends are met. First, the label is not from the perspective of an adherent, thus affording a new language of analysis. Second, the label describes two essential features of the model; that it is (1) based on changes in recall produced by varying qualitative aspects of the stimuli (e.g., phonological similarity, word-length, and articulatory suppression) and that it is (2) described in structural terms in the literature developing theory from those qualitative changes.

Chapter 2 presents relevant theory and research in order to outline the strengths and flaws of this qualitative/structural model. Briefly, this model provides a good account of many areas of research in the memory domain. However, the main deficiencies of the qualitative/structural model of working memory are its description of working memory as a set of structures and its reliance on recall as the only measure of working memory's products instead of recall and order information.

Theoretical definition of working memory as a set of structures arose from data gathered using the qualitative and structurally based paradigm. As reviewed earlier, Baddeley and Hitch (1974) observed that a secondary digit span task was performed as well when a primary task, also presumably requiring the STS, was concurrently performed as when either the primary or secondary task alone was performed. Baddeley

and Hitch suggested that there may be multiple short term stores that were in some way related and controlled by a control process, a working memory. Subsequent investigation led to working memory being described as comprising of a Central Executive responsible for task distribution, attention, and temporary storage (Baddeley, 1966c, 1984, 1986, 1990, 1992a, 1992b; Baddeley, Logie, Bressi, Della Salla, & Spinnler, 1986; Baddeley, Bressi, Logie, Della Salla, & Spinnler, 1991; Gathercole & Baddeley, 1993; Morris & Baddeley, 1988; Morris, Craik, & Gick, 1990; Morris & Jones, 1990b; Reisberg, Rappaport, & O'Shaugnessey, 1984; Spinnler, Della Sala, Bandera, & Baddeley, 1988; Van der Linden, Coyette, & Seron, 1992), a Phonological Loop (Baddeley, 1966a, 1966b, 1984, 1986, 1992a, 1992b; Baddeley, Lewis, & Vallar, 1984; Baddeley, Thompson, & Buchanan, 1975; Ellis & Hannelley, 1980; Longoni, Richardson, & Aiello, 1993; Salame & Baddeley, 1982; Vallar & Baddeley, 1984), and a Visuo-Spatial Sketch-Pad (Baddeley, 1986, 1992a, 1992b; Farah, Hammond, Levine, & Calvanio, 1988; Hanley, Young, & Pearson, 1991; Logie, 1986; Morris, 1989).

AN OVERVIEW OF BADDELEY'S DEFINITION OF WORKING MEMORY

In describing working memory, Baddeley's (1986, 1992a, 1992b) most current reviews of his work will be used in order to accurately present how he and his colleagues currently conceptualize the construct of working memory (Figure 1).

Currently, working memory is described as comprising of a Central Executive and two 'slave' systems (Baddeley, 1992a, 1992b). The slave systems are a Phonological Loop which is responsible primarily for verbal material and a Visuo-Spatial Sketch-Pad responsible for visual and spatial location memory (Figure 1). With two systems, there was a perceived need for some type of controller to monitor those two systems, an idea suggested in the work of Atkinson and Shiffrin (1968, 1971). Baddeley proposes that a 'Central Executive' functions as the construct explaining how information is allocated to a slave

system, and how the output of the slave systems is monitored and integrated.

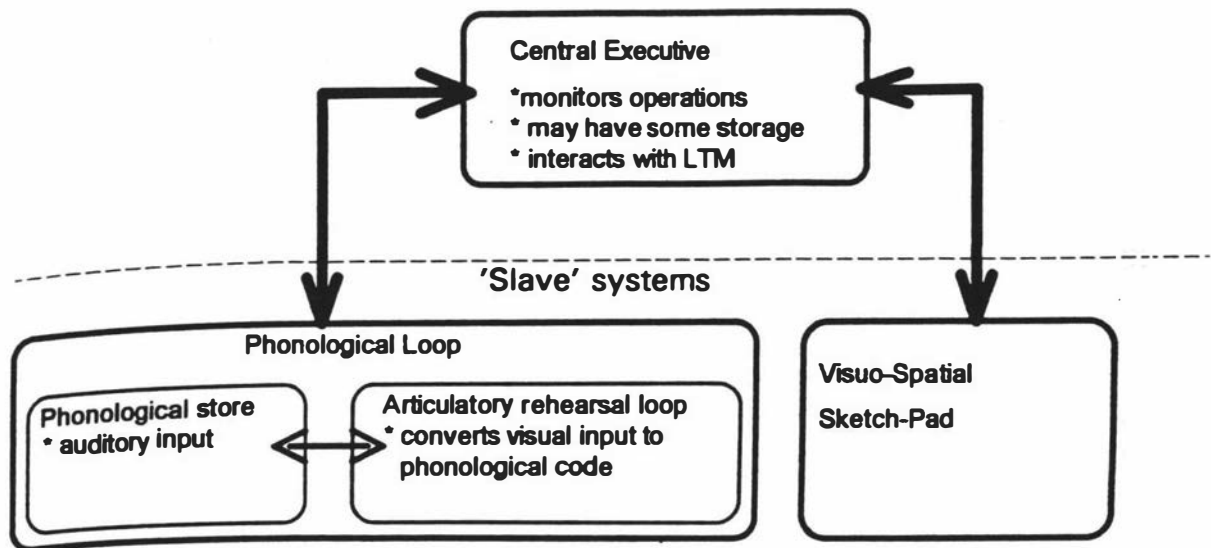


Figure 1. A model of working memory showing the Central Executive, the Phonological Loop, and the Visuo-Spatial Sketch-Pad.

Baddeley's (1986, 1992a) qualitative/structural model represents structures (more than processes; A.D. Baddeley, personal communication, April 19, 1993) that appear to operate relatively separately from each other but that are also interdependent. The components of the qualitative/structural model have been derived primarily from data where two tasks are compared in the effect they have on the output from the person. The logic underlying a qualitative and structurally based paradigm is that if two comparison tasks use the same component, then performance will decrease if resources are limited for that component. If two tasks use two distinct components, then little performance change should be observed from using the two components alone to when the tasks are performed concurrently.

THE QUALITATIVE/STRUCTURAL MODEL OF WORKING MEMORY

This first section presents an overview of the evidence both for and against the positing of a separate Phonological Store

and Visuo-Spatial Sketch-Pad, both of which are monitored by a Central Executive.

The Phonological Loop

The Phonological Loop is currently hypothesized as consisting of a phonological memory store of about 1.5 to 2.0 seconds duration and an articulatory control system which both converts non-phonological material into phonological material and which allows sub-vocal rehearsal (Baddeley, 1986, 1992a, 1992b). The evidence cited in support of the Phonological Loop includes the phonological similarity effect, the irrelevant speech effect, the word length effect, neuropsychological evidence from persons with an impaired STS, and performance of persons with dysarthria (Baddeley, 1986, 1992a, 1992b).

The Phonological Similarity Effect

The phonological similarity effect is where more phonologically dissimilar items are recalled than phonologically similar items. This effect has been replicated many times with the accepted conclusion being that the phonological similarity effect provides evidence that the Phonological Loop is based on an articulatory code (e.g., Baddeley, 1966a, 1966b; Coltheart, 1993; Conrad, 1964; Conrad & Hull, 1964; Ellis, 1980; Henry, 1991; Hulme, & Tordoff, 1989; Longoni, et al., 1993; Richardson, Greaves, & Smith, 1980; Schweikert, Guentert, & Hersberger, 1990). That increased phonological similarity produces lower recall of items is also consistent with an interpretation that words or letters are 'held' in working memory phonetically, as described in the span model of the STS (Atkinson, & Shiffrin, 1968, 1971). Thus, the qualitative/structural model of working memory continues with a phonologically-based store as proposed in the span model.

This Phonological Store is presumed to receive information directly when stimuli are presented auditorily and indirectly when presented visually (Baddeley, 1986, 1992a, 1992b). These

inferences have been made based on the discriminative effect of articulatory suppression, a rapid and continuous vocalisation of some very well learned vocal sounds (e.g., repeatedly saying the or counting aloud from 1 to 8 repeatedly). Articulatory suppression is presumed to block covert subvocalisation (Murray, 1967, 1968). It has been generally agreed that preventing subvocalisation also prevents subvocal rehearsal and, by implication, the rehearsal of items by the Articulatory Rehearsal Process (Baddeley, 1986, 1992a, 1992b; Besner, 1987; Murray, Rowan, & Smith, 1988). When stimuli are presented auditorily, the phonological similarity effect remains with or without articulatory suppression (Baddeley, et al., 1984; Longoni, et al., 1993; Salame & Baddeley, 1982). In contrast, it has been observed that the phonological similarity effect is absent when articulatory suppression is in use and stimuli are presented visually (Baddeley, et al., 1984; Besner & Davelaar, 1982). The received interpretation of these results is that with auditory presentation, stimuli have direct access to the Phonological Store, and thus show a phonological similarity effect. When stimuli are presented visually, they are first routed through an Articulatory Rehearsal Process which converts visual to articulatory codes and stores them in the Phonological Store. Provided the stimuli are recoded into articulatory codes (as when articulatory suppression is absent), a phonological similarity effect is again present. When articulatory suppression is used, the Articulatory Rehearsal Process is prevented from converting visual to articulatory codes, and hence the stimuli do not enter the Phonological Store. Thus, because the stimuli are not in the Phonological Store, there can be no phonological similarity effect (Baddeley, 1986, 1992a, 1992b; Besner, 1987)¹.

¹ It could be argued that, if articulatory suppression blocks rehearsal, almost no recall ought to occur. In fact, some recall does remain with articulatory suppression. This anomaly has been explained by inferring that the Central Executive has some 'storage' ability in addition to the storage provided by the Phonological Store (Baddeley, 1986), and may use LTM as a temporary storage area (A.D. Baddeley, April 19, 1993, personal communication).

Flaws with a single Phonological Store account of the phonological similarity effect

While the above interpretation of the phonological similarity effect as an indicator of a Phonological Store is appealing, there are two major aspects that this account cannot explain. First, it has been shown that articulatory suppression does not necessarily impair a person's ability to make decisions based on the phonological characteristics of items. Second, where the qualitative/structural model of working memory assumes that item and order information are one, there are data which suggest otherwise.

Articulatory suppression does not suppress all phonological codes

Besner and Davelaar (1982) presented short lists of letters to their participants to recall in serial order at the end of the list. There was a phonological similarity effect when articulatory suppression was not used which was eliminated with articulatory suppression. Thus Besner and Davelaar completely replicated the data upon which Baddeley's (1986) description of the Phonological Loop was based. However, there were two additional conditions in the Besner and Davelaar study: a list of pseudohomophones (e.g., Brane, Skule) and a list of non-words (e.g., Frane, Zule). The logic was that "if suppression does not prevent lexical access from print via some form of phonological code, then letter strings which sound like real words should be better recalled than letter strings which do not" (p.703). Stated the other way around, if articulatory suppression prevents all phonological coding, then there would be no way that persons could use the sound of the pseudohomophones to assist their later recall and obtain better recall of pseudohomophones over nonwords. However, the data showed that, even with articulatory suppression, the identical sound of the pseudohomophones to their correctly spelled derivatives provided an advantage in later recall over the nonwords. The conclusion was that articulatory suppression only prevents some types of phonological coding. However, the wider implication is that, if the same logic used to propose Baddeley's (1986, 1992a,

1992b) original Phonological Store is used upon these data, one could reasonably propose a second Phonological Store, as was done (Besner & Davelaar, 1982). This issue is discussed further in Experiment 5 of the present dissertation.

Order information has not been separated from item information

The second flaw with the present account of the Phonological Store of working memory is that preservation of order information has not been accounted for adequately. There is some empirical work that suggests order information is important in the production of acoustic similarity effects in the STS (e.g., Conrad, 1964; Gruneberg & Melton, 1972; Healy, 1974, 1982; Lewandowsky & Murdock, 1989; Watkins, et al., 1974; Wickelgren, 1965). Because (1) acoustic similarity or phonological similarity effects are used to define the Phonological Loop and (2) because there is an implied relationship between the STS and working memory, then there would appear to be a need for the qualitative/structural model of working memory to specify how order information is maintained. Instead, a focus on content has been maintained in most literature on working memory, and as will be discussed throughout this dissertation, has been to the detriment of models of working memory. As Burgess and Hitch (1992) write in discussing the usefulness of Baddeley and Hitch's (1974) model of working memory: "...given that one of its [the Phonological Loop's] major functions is the preservation of order information, surprisingly little is said about how this is achieved, and some of what is said is clearly incorrect." (p.431).

While the reasons for failing to examine order information explicitly are unclear, there do exist some data which suggest what the role of order information might be in relation to working memory and the phonological similarity effect. For example, the Conrad and Hull (1964) study (which is quoted in much of the literature on the phonological similarity effect) did not simply examine whether more phonologically dissimilar than phonologically similar letters were recalled, rather, they also compared the effect of

repeated selection from a 3-item or from a 9-item pool in construction of 6-letter sequences. Conrad and Hull (1964) cautioned that they would "hesitate on the present data alone to claim a simple relationship between span and [acoustic] confusability. There are inevitably structural differences between the 3 and the 9-letter vocabularies. In particular, more letters are repeated in the sequences drawn from the smaller set." (p.431). In this statement is contained a hint that the phonological similarity effect is also related to the repetition of items and hence may be an effect of either item or order confusion. This inference is supported by data from Richardson (1984) which showed smaller phonological similarity effects using a free-recall paradigm than in a serial recall paradigm. From this differential effect on the phonological similarity effect, Richardson (1984) also suggested that the Phonological Store contributes only to those tasks in which accurate serial-order information was required.

Coltheart (1993) has presented two studies directly bearing upon the issue of whether the phonological similarity effect is due to content and/or order confusions. At the time that the present author had just completed Experiment 6 of the present dissertation, Coltheart published an independent investigation of the question asked in my Experiment 5. Coltheart presented subjects with a series of either phonologically similar or phonologically dissimilar words. Typically a phonological similarity effect would be found in such a paradigm. Coltheart also constructed her word lists from either a fixed pool of 10 stimuli or presented novel words on each trial. The repeated list represented a condition in which the stimuli would be easily recalled but the order would be changed on each trial. The novel list represented a condition in which both the stimuli and the order would need to be recalled. The logic of this manipulation was that if the phonological similarity effect was an effect of order, then the repeated list (order only) would show a larger phonological similarity effect than the

novel list (content plus order). Coltheart's results did not find any difference between repeated or novel lists. However, as will be discussed in the introduction to the present Experiment 5, her results consistently showed a non-significant (at $p < .05$) tendency toward a larger phonological similarity effect for repeated than for novel lists. The relationship between order information and the phonological similarity effect is clearly an area for further investigation.

Summary

In summary, there are some empirical reasons why it cannot be assumed that the phonological similarity effect is definitive of a Phonological Store in the exact form proposed by Baddeley (1986, 1992a, 1992b). In some conditions phonological information clearly remains available when the Phonological Store, the presumed repository of phonological information, is considered inoperative. Also, there is some evidence suggesting that the Phonological Store is as concerned with order as with content information. Neither of these results can be adequately explained by the current qualitative/structural model of working memory.

The Irrelevant Speech Effect

The irrelevant speech effect is when a noise extraneous to the recall task reduces the level of recall of visually presented materials. The effect has been replicated many times (Colle, 1980; Colle & Welsh, 1976; Jones & Macken, 1993, in press; Jones, Madden, & Miles, 1992; Jones, Miles & Page, 1990; Miles, Jones, & Madden, 1991; Morris & Jones, 1990a, 1990b; Morris, Jones, & Quayle, 1989; Salame & Baddeley, 1982, 1986a, 1986b, 1987, 1989, 1990). Irrelevant speech disruption is independent of the semantic characteristics of the speech, and is now thought to be due to the change in composition from one sound to the next in the irrelevant stream. This account of the irrelevant speech effect is termed the *changing state hypothesis* (Jones, & Macken, 1993, in press; Jones, et al., 1992).

Baddeley (1992a) assumes "that irrelevant speech gains obligatory access to the phonological store and is thus able to corrupt the memory trace, leading to impaired recall" (p.9). Based on this interpretation, the irrelevant speech effect has been perceived as supporting the Baddeley (1986, 1992a, 1992b) version of the Phonological Store in two ways. First, the irrelevant speech effect has been interpreted as demonstrating that the Phonological Store is indeed acoustically based, as auditory information appears to take precedence over visual information. Second, the irrelevant speech effect has been interpreted as demonstrating that auditory material has obligatory access to a Phonological Store (Baddeley, et al., 1984; Salame & Baddeley, 1982, 1987, 1989).

Baddeley's (1986) inferences from the irrelevant speech effect have not been without criticism. The inference that the irrelevant speech effect has obligatory access at the phonological level (Salame & Baddeley, 1982, 1989) has recently been challenged by Jones and colleagues, who are now presenting a more process-oriented model of how the irrelevant speech effect operates (Jones, 1992; Jones & Macken, in press). In its simplest form, Jones and Macken (in press) suggest that, when successive items are presented, there is a set of cues to the serial order of those items established in the form of cognitive linkages between those items (or objects). The linkages are held to be more robust when derived from items with a high *changing state value* (e.g., phonologically dissimilar items) than when they have a low *changing state value* (e.g., phonologically similar items). When irrelevant items with a high *changing state value* compete with to be recalled items in serial order, there is disruption of the order cues. When irrelevant items with a low *changing state value* compete with serial order information, there is less disruption on recall. Thus, as with the phonological similarity effect, there is an alternative interpretation of the irrelevant speech effect that, unlike the qualitative/structural model of working

memory, bases its mode of operation upon disruption of item order in opposition to item content.

In summary, the irrelevant speech effect has been presented as evidence for the Phonological Store. Again, there is also evidence that suggests that the irrelevant speech effect is as much involved with item order as with content. In part, the difference between the two perspectives is again related to whether working memory is described as a set of structures or a process.

The Word-Length Effect

The word-length effect is when fewer long words are recalled than shorter words. Word length is not determined solely by syllable length nor the number of letters. Instead, word-length effects are best predicted when word length has been determined from the time it takes to articulate (read aloud) those words. Consideration of the word-length effect allows further development of the Phonological Store.

Evidence presented in support of an Articulatory Rehearsal Process based on the word-length effect

Baddeley, et al. (1975) conducted a series of eight experiments examining the word-length effect. Their basic procedure was that lists of short or long words were presented to participants and then participants were required to recall the items in serial order. In their first experiment, 1- and 5-syllable words were presented to participants in an incrementing span task. Over presentation spans of from 4 to 8 items, 1-syllable words were always better recalled than 5-syllable words. At this point there could be many alternate explanations of this word-length effect based on for example word frequency or articulation rate. So, in their third experiment, Baddeley, et al. constructed two lists of di-syllabic words matched for frequency, but with items in one list taking longer to articulate than in the other. When these two lists were presented for later serial recall, the list that took longer to articulate was less well recalled than the list that took

less time to articulate. In their fourth experiment, when the lists were also matched for the number of phonemes, frequency, and syllables, there was still a word-length effect. From these data, the authors concluded that the word-length effect was related to the articulation time of the items. When articulatory suppression was present, Baddeley, et al. observed that the word-length effect disappeared with visual but not with auditory presentation (see also, Levy, 1971; Longoni, et al., 1993; Peterson & Johnson, 1971). However, Baddeley, et al. (1984) found data which suggest that, provided articulatory suppression was present during acquisition and recall, there was a word-length effect irrespective of presentation modality.

Word-length effects found by Baddeley, et al. (1975) have also been used as a basis from which to infer the duration of an item in working memory. The duration of working memory was calculated as the number of words recalled divided by the articulation time of those words. For example, if a person could recall 5 words and could articulate words of that type at 2.5 words per second, then the duration of working memory would be calculated as $5/2.5 = 2$ s. Empirical investigation of the duration of the Phonological Store using the above logic found that working memory duration ranges across persons from 1.5 to 2.0 s (Baddeley, et al., 1975).

The word-length effect has proven to be a replicable result cross-culturally (e.g., Ellis & Hannelley, 1980; Naveh-Benjamin & Ayres, 1986). Ellis and Hannelley's (1980) data show that native Welsh speakers speaking Welsh digits had a shorter span on the Welsh than the English language version of the WISC. However, when the total time to articulate the digits was measured, no difference between English and Welsh speakers was observed. The data from Ellis and Hannelley suggests that the Phonological Store is limited by temporal duration, rather than by item capacity per se.

In summary, the word-length effect appeared to be related to the inverse of the time to articulate items, articulatory

suppression prevents the word-length effect, and the duration of working memory was calculated at 1.5 to 2.0 seconds.

To account for these results based on the word-length effect, Baddeley and his colleagues (Baddeley, 1986, 1992a, 1992b; Baddeley, et al., 1984; Baddeley, et al., 1975) have proposed an *Articulatory Rehearsal Process* which will now be described. The rehearsal speed of the *Articulatory Rehearsal Process* is presumed to be determined by the rate at which the items can be articulated. Hence, the more rapidly items can be articulated, the more rapidly each item can be rehearsed. The more rapidly an item can be rehearsed, the more time is left (out of the total working memory duration of 1.5 to 2.0 seconds) for other items (in the *Phonological Store*) to be rehearsed before they fade below a threshold below which they are forgotten. Consequently, when recall for a set of items which can be rehearsed rapidly (e.g., short words) is compared to recall for a set of items which are less rapidly rehearsed (e.g., long words), recall will be greater for the set with the faster articulation rate per item. Thus, the *Articulatory Rehearsal Process* accounts for the production of word-length effects and also highlights that it is rate of rehearsal, and not solely capacity of storage, that determines how effective working memory is (Baddeley, 1986).

Thus the *Phonological Loop* is not only comprised of a *Phonological Store*, but also is presumed to have an *Articulatory Rehearsal Process* (Figure 1). The *Articulatory Rehearsal Process* maintains items and converts visual to articulatory information. The *Articulatory Rehearsal Process*, but not the *Phonological Store*, is vulnerable to disruption by concurrent articulation and irrelevant speech. Thus, the *Articulatory Rehearsal Process* and the *Phonological Store* are seen as two separate but complementary processes (Longoni, et al., 1993).

Flaws with an Articulatory Rehearsal Process based on the word-length effect

The model of the *Articulatory Rehearsal Process* stands or falls on the relationship between articulation rate (the mean rate of reading the set of stimulus items aloud) and recall.

Any empirical criticism of this theory must demonstrate that articulation rate and recall are either not related or are mediated via a third cognitive process. On both issues, there is some empirical work contraverting the relationship between articulation rate and recall.

Tehan and Humphreys (1988) presented participants with high and low frequency words varying across three word classes (adjective, noun, function). Measures of reading rate and pronunciation rate were also taken. Word frequency effects could be predicted by articulation rate differences between high and low frequency items. However, articulation rate differences across word class failed to produce word class effects in the predicted direction. From Tehan and Humphrey's results, the data suggest that *span differences can occur without pronunciation rate differences*. Furthermore, span differences were as marked with, as without, articulatory suppression, which ought to have prevented pronunciation rate differences having an impact on later recall.

Caplan, Rochon, and Waters (1992) also present data in which sets of words were varied across the time to articulate those sets. When these sets of words were also matched for the number of syllables and phonemes in them, there were no word-length effects. From these data, Caplan, et al. concluded that it was the phonological structure of the words, not the articulation time, that produced the word-length effect. Baddeley and Andrade (1994) countered that the sets used by Caplan, Rochon, and Waters were not sufficiently different in articulation time and that the word-length effect was in fact a robust phenomenon that has been replicated many times, hence must be disconfirmed repeatedly.

In their reply to Baddeley and Andrade (1994), Caplan and Waters (1994) suggest that in some of these 'robust' studies (e.g., Ellis & Hannelley, 1980; Hoosain & Salili, 1988), there was a potential confound of phonological similarity with articulatory duration. That is, in the cross-cultural studies, there is no indication of phonological similarity having been controlled (c.f., Longoni, et al., 1993). Also,

Caplan and Waters suggest that measuring articulation rate on different subjects (Baddeley & Andrade) and only on a sample of the total set of target words is an insufficiently precise measurement of articulation rate. They suggest that *all items* from the set be presented to the *same subjects* as those who performed the recall task. When these methodological refinements were enacted, there was again no correlation found between articulatory rehearsal rate and recall (Caplan & Waters). Thus, from the work of Caplan and colleagues, the suggestion is that the typical measurement of articulation rate is imprecise, that phonological similarity may have confounded previous replications of the relationship between word-length effects and articulation rate, and that, as a result, *word-length effects cannot be described as unambiguously produced by differing articulation rates.*

There has also been evidence that the word-length effect is not produced solely during rehearsal of items in a single input buffer, but that it is an effect produced at recall by the operation of an output buffer. Cowan, Day, Saults, Keller, Johnson, and Flores (1992) performed a detailed analysis of the relationship between recall time and word-length effects. Their data showed that the longer an item took to recall, the lower was its probability of recall. That is, *word-length effects might be produced by decay at output*, rather than as a function of active rehearsal of items by the Articulatory Rehearsal Process.

In an extension of the work by Cowan, et al. (1992), Avons, Wright, and Pammer (1994) compared serial recall and probed recall for word-length effects in an adult population. Their reasoning was that if the Articulatory Rehearsal Process operates as theorized (Baddeley, 1986, 1992a, 1992b), then there should be no difference in the magnitude of the word-length effect as a function of recall method. However, if the word-length effect were partly occurring because it takes more time to prepare an articulatory code for longer over shorter words, with that delay having a deleterious effect on an output buffer, then probed recall should show less of a

word-length effect than serial recall. That is, if the word-length effect is the result of a separate output buffer for speech (Coltheart, Avons, & Trollope, 1990; Monsell, 1987), then there would be greater word-length effects for serial recall (examining the output buffer) than for probed recall (examining the input buffer). This is exactly what the data in the studies by Avons, et al. show.

In addition, in Avons, et al.'s (1994) first two experiments, a five second delay before recall was imposed between the trial and recall. To maintain items over this period persons would be expected to engage in rehearsal. If rehearsal is the source of the word-length effect, then the magnitude of that effect ought to have become larger for delayed than for immediate recall. Avons, et al.'s results showed that the word-length effect continued to be different as a function of recall method even after a delay. Of most importance, delayed recall did not increase the word-length effect, suggesting that *additional rehearsal does not necessarily produce increased word-length effects*, a conclusion at odds with the received interpretation of how the Articulatory Rehearsal Process operates. From their results, Avons, et al. concluded that *"...some, if not all, word-length effects arise at output, possibly by decay during output, by restrictions imposed by a limited-capacity output buffer, or by output interference."* (p.229).

Henry (1991) also compared word-length effects in 5 and 7-year-old children. The importance of the two ages is that 7, but not 5-year-old children are thought to use rehearsal as a memory strategy. Therefore, if the word-length effect occurs as a function of rehearsal, then 7 and not 5-year-old children would display word-length effects. Word-length effects in both serial and probed recall were shown by 7-year-old children. However, 5-year-old children also showed word-length effects for serial recall. Such a result is inconsistent with the word-length effect being produced by a rehearsal process, because children are not thought to be using rehearsal at that age. Thus, some evidence exists that

word-length effects can occur when rehearsal is not presumed to be operating. It must be noted that such a conclusion relies on the assumption that 5-year-old children do not use rehearsal, an assumption that definitely requires greater empirical scrutiny.

Thus, there do exist some data that show that (1) the word-length effect occurs in the absence of articulation rate differences, (2) that the word-length effect is related to the time taken to recall items, (3) that increased rehearsal time does not necessarily increase the word-length effect, and (4) that word-length effects may occur in the absence of rehearsal. When considered as a body of evidence, these results suggest that the word-length effect is not necessarily produced by different rates of rehearsal as originally suggested. Instead, the word-length effect might be better considered as resulting from the operation of a separate output buffer for speech (Avons, et al., 1994; Coltheart, et al., 1990). This proposed separation of rehearsal (at input) and speech output processes represents a critical theoretical issue for the qualitative/structural model of working memory which is based on a unitary Articulatory Rehearsal Process.

Summary

In conclusion, the word-length effect has been used to support a dual-process model of working memory. The word-length effect has been taken to indicate that a separate process converts visual to articulatory information and rehearses that information from a Phonological Store. This interpretation has endured and been supported by many replications of the word-length effect. However, recently there has been evidence advanced that the relationship between articulation rate and word-length effects is not always robust and that the word-length effect is possibly an effect produced by a speech output buffer (Avons, et al., 1994; Caplan, et al., 1992; Caplan & Waters, 1994; Coltheart, et al., 1990; Cowan, 1992; Cowan, et al., 1992; Henry, 1991; Monsell, 1987).

Neuropsychological Evidence

Turning to neuropsychological evidence, Baddeley (1986, 1992a, 1992b) asserts that persons with STS deficits have in fact a defective Phonological Store (e.g., Baddeley, Vallar, & Wilson, 1987). The operation of the Phonological Store is hypothesized to produce the phonological similarity and word-length effects. Thus to the extent that persons can recall words presented visually, if it is the Phonological Store that is damaged, then both the phonological similarity and word-length effects should be absent. This is what has been found, thus supporting a close link between the STS and working memory (Vallar & Baddeley, 1984). However, there is also more recent evidence that concludes that the STS and working memory are in fact separate (e.g., Cantor, Engle, & Hamilton, 1991; Klapp, Marshburn, & Lester, 1983; Swanson, 1993, 1994). Thus, the statement that the STS deficits are Phonological Store deficits remains unresolved.

Evidence from Persons with Dysarthria

Finally, persons who are dysarthriac (mute) have been shown to develop both phonological similarity and word length effects by measuring their written recall of words and letters (Baddeley & Wilson, 1985). Also, Bishop and Robson (1989) have found phonological similarity and word-length effects in persons who have never had speech. If the locus of the word-length effect was at a purely speech level, this effect would be incongruent with Baddeley's (1986) model of the Phonological Loop. Thus it would seem that the Articulatory Rehearsal Process is not dependent on the capacity to speak. As Baddeley (1992a) suggests, the Articulatory Rehearsal Process may be based on some more central articulatory process than speech output alone.

Summary

To summarize, the Phonological Loop has been hypothesized as a distinct process that is primarily phonological, is temporally limited to about 2.0 seconds, uses subvocal

rehearsal to maintain existing phonemes, and that is capable of converting other modalities to a phonological mode.

While this model of the Phonological Loop has the appeal of parsimony and also accounts for many empirical data, there are also some data which cast doubt on the model. Specifically, the present version of the Phonological Loop fails to provide an adequate explanation of the separate processing of item and order information, and the word-length effect may be as much an effect at output as at rehearsal.

I shall now provide a brief outline of the Visuo-Spatial Sketch-Pad and the Central Executive. As these two components are not explicitly investigated in the present dissertation, they will not be reviewed in as much detail as the Phonological Loop.

The Visuo-Spatial Sketch-Pad

The second 'slave' system, the Visuo-Spatial Sketch-Pad, is a necessary second system to explain how words presented visually can be recalled when articulatory suppression prevents subvocal rehearsal operating and the Phonological Store being used. Baddeley (1992a) describes a task in which a person remembered a sentence of instructions by either visualizing the sentence or by rote memory (the Phonological Loop). For example, a spatial sentence might be:

" In the starting square put a 1, in the next square to the right put a 2, in the next square down put a 3, etc"
(Baddeley, 1992a, p.10).

A second parallel sentence was presented in which the spatial adjectives were replaced by non-spatial adjectives (e.g., 'right' might be replaced with 'good'). When persons tracked a moving spot of light (a visuo-spatial task), recall of the sentences with spatial adjectives decreased, but not recall of non-spatial adjective sentences. When persons then tracked a moving pendulum while blindfolded (a spatial-only task), greater recall decrements for spatial sentences occurred than in the visuo-spatial tracking task (Baddeley & Lieberman,

1980). From this Baddeley concluded that the Visuo-Spatial Sketch-Pad is a distinct component and that it is perhaps more spatial than visual in nature.

However, subsequent studies have shown that there is also a definite visual component, in addition to the spatial component of the Visuo-Spatial Sketch-Pad (Farah, et al., 1988; Hanley, et al., 1991; Logie, 1986) .

Thus, it appears that the Visuo-Spatial Sketch-Pad consists of processes that temporarily store information in visual or spatial form. As yet, the duration of the storage and how the visual and spatial components are interrelated has not been explicated. Neither has a process analogous to subvocal rehearsal in the Phonological Loop been found in the Visuo-Spatial Sketch-Pad. Finally, there appears to have been less attention paid to clinical deficits in the Visuo-Spatial Sketch-Pad than in the Phonological Loop.

The Central Executive

Baddeley (1992a) refers to the Central Executive as an *"area of residual ignorance rather than a well-worked-out concept"* (p.12). Yet, where there are at least two working memory components (the Phonological Loop and the Visuo-Spatial Sketch-Pad), it is intuitively sensible to assume that some mechanism is responsible for controlling and monitoring the output of these systems (e.g., Atkinson & Shiffrin, 1971). Furthermore, as processing is ongoing with considerable dropping of irrelevant information (Waldrop, 1987), perhaps it is also the Central Executive that selects what is relevant for potential processing. This selection process is remarkably similar to an attentional mechanism, a point Baddeley (1986, 1992a, 1992b) has noted.

In comparing the Central Executive to an attentional model, Baddeley (1986, 1992a) uses a model of attention proposed by Norman and Shallice (1980, cited in Baddeley, 1992a). In the Central Executive, it was hypothesized that processing is typically controlled by activating a schema and periodically monitoring the output. Where the output fails to meet the

desired criteria, the Central Executive interrupts the current schema, selects a new strategy, and activates that strategy. From this model of the Central Executive, Baddeley (1986) proposed an explanation for the results of a study by him twenty years earlier (Baddeley, 1966c), in which he examined the ability of persons to generate random letter sequences. Baddeley's (1986) explanation of his earlier results was as follows. Letters are learned in an alphabetical sequence a, b, c, d, etc, termed an alphabetical schema. To produce a random sequence of letters, the alphabetical schema must be activated and monitored for one letter (x_i). Next, a decision rule must be invoked to ensure that the subsequent letter (x_{i+1}) is not predictable by a series function ($x_{i+1} \neq f(x_i)$) and that the subsequent letter has not been repeated more frequently than any other previous letter ($p(x_{i+1}) = p(x_i: i = 1 \text{ to } i)$). When these conditions are satisfied, then a new production schema can be activated. If the above explanation is an accurate portrayal of Central Executive processes, when the Central Executive is overloaded (assuming it has limited capacity) redundancy of letters and stereotypy of sequence would occur. This is the exact effect Baddeley (1966c) had obtained early in his career in what he, at the time, considered a puzzling result. Thus it would appear that the Central Executive monitors output and selects strategies or schemata for processing.

A second source of information about the Central Executive comes from the performance of persons with Alzheimer's Dementia. Morris and Baddeley (1988) compared the performance of persons on either a task previously hypothesized as using the Phonological Loop or a task using the Visuo-Spatial Sketch-Pad against performance using both slave systems at once. Using either slave system alone should only require one activation of that system by the Central Executive. Using both systems concurrently was assumed to require activation of the Central Executive every time a switch between systems was made. Because the Central Executive model was based upon models of attention, because attention is perceived as a

frontal lobe process, and because post-mortum examinations suggest Alzheimer's patients have frontal lobe atrophy, persons with Alzheimer's Dementia were predicted to have a deficit in the Central Executive so that they would be less able to perform the necessary switching between systems. The prediction was that persons with Alzheimer's Dementia would show a significantly greater decrement in recall when performing both tasks than when performing one task alone. Persons of the same age as the Alzheimer's Dementia group and younger people were predicted to show no significantly larger decrement in recall when both systems were in use than when either system alone was in use. These are precisely the results that Morris and Baddeley obtained. These results suggest that the concept of the Central Executive is modelled correctly as a monitor and task assigner and is a potentially useful indicator of Alzheimer's Dementia and possibly other clinical executive deficits.

To summarize, the Central Executive is considered by Baddeley and colleagues as a necessary theoretical link between the Phonological Loop and the Visuo-Spatial Sketch-Pad. There is also increasing empirical support for the concept of the Central Executive as a manager of processing. The management of processing at this stage appears to involve monitoring of processing in addition to activation and inhibition of processing. There is also some data that suggest that the Central Executive has some storage ability of its own (Morris, 1989), and that it can temporarily create a type of memory cache of its own (Reisberg, et al., 1984).

SUMMARY

In this dissertation, how the qualitative/structural model is able to account for the storage of visually presented verbal information is being investigated. The particular explication used for this model was that presented by Baddeley (1986). The qualitative/structural dual-task model of working memory is able to explain many data generated experimentally. However, it is also the case that sometimes data do not

readily fit this working memory model (e.g., Klapp, et al., 1983). Specifically, the above review highlighted that the qualitative/structural model does not provide an adequate model of how verbal order information is preserved, and second, there has been some evidence suggesting that the word-length effect is an effect at output, and thus not directly part of what is typically thought of as the Articulatory Rehearsal Process of working memory. Thus, there is a need to examine the role of order information and the operation of the word-length effect, in addition to the already stated intention of investigating similarities and differences in how the qualitative/structural and quantitative/process models of working memory account for storage of visually presented verbal information.

To foreshadow this investigation it must be noted that, because the present work focuses on verbal information, only tasks presumed to index the Phonological Loop of the qualitative/structural model will be used. This is also because the quantitative/process model presented below only deals with verbal information. Thus, by focussing on only verbal information, one is comparing 'apples with apples'. However, this does not imply that Baddeley's (1986) model of working memory can be reduced to the Phonological Loop alone.

CHAPTER 3: QUANTITATIVE/PROCESS MODELS OF WORKING MEMORY

Evolving from the development of the qualitative/structural model of working memory, two other applied cognition areas were also making use of the concept of working memory: the area of reading (e.g., Daneman & Carpenter, 1980, 1983) and the area of aging (e.g., Salthouse, et al., 1989). Both of these areas define the construct of working memory as consisting of an interaction between operation and span components. That is, both of these areas appear to have used the Baddeley and Hitch (1974) method of *measuring* working memory as both an operational and conceptual definition. In the present dissertation, an operation plus span definition of working memory will be termed a *complex-span* task at an operational level to help alleviate some of the apparent overlap between the concept of working memory and tasks used to measure that concept. In contrast, measurement of span only will be termed a *simple-span* task.

The utility of a complex-span measure of working memory was that it allowed researchers to more accurately predict, for example, reading comprehension performance (Daneman & Carpenter, 1980) using a complex-span task than could be done with a simple-span task. Because complex-span methods of measuring working memory obtain a score or quantity that represents the *span* of working memory and because they refer to working memory as an *interactive process*, this group of working memory models will be referred to as quantitative/process models in this dissertation. Thus, complex-span measures are used to infer quantitative/process models of working memory.

Before proceeding, it is important to consider whether the similarity of approaches described below can be inferred to comprise a model. Most apparently, the approaches used below are in no way as clearly developed as Baddeley's (1986) model. Second, while the approaches described below do share a common theme, as will be developed, they are not identical.

And so, in the present context, it is important to be aware that using the term 'model' for both quantitative/process and qualitative/structural models of working memory does not imply equality of either development or theoretical scope between them.

Generally, quantitative/process approaches to operationalising working memory are based on psychometric principles of reliability and validity (Anastasi, 1982; Baddeley, 1986). These principles imply that working memory differences are reflected in score differences. That is, differences on a score (or quantity) reflect or imply differences in working memory (processes). This very concise operational definition allows working memory to be examined in relationship to other processes. The quantitative/process model of working memory relies on the complex-span measure of working memory being able to describe and predict empirical data. It is assumed that if a measure is able to describe and predict data, that the underlying principles in constructing the measure are theoretically relevant (Anastasi, 1982).

This process of defining working memory suggests a method for examining the application of the quantitative/process theory of working memory. First, Chapter 3 will examine how working memory has been operationalised and how successful this operationalisation has been. The conclusion of this first section is that the operation plus span method of defining working memory has been both reliable and valid in psychometric terms. Having established the pragmatic use of complex-span tasks, the second section examines what the generic complex-span measure implies about the structure and operation of quantitative/process model of working memory. The conclusion of this section is that the operation plus span measure involves preservation of content, order, and perhaps processing speed.

A PRAGMATIC ANALYSIS OF THE COMPLEX-SPAN TASKS: DO THEY PERFORM?

In examining the success or otherwise of the complex-span operationalisation of working memory, the succeeding sections

discuss the definition of a complex-span task, the reliability of operationalisations of a complex-span task, and the validity of a complex-span task as a measure of a quantitative/process model of working memory.

Table 2. A selective review of working memory definitions used by researchers using a quantitative and process oriented correlational approach to aging and working memory (italics added) .

Campbell & Charness (1990)	"... the cognitive processes used to manipulate or temporarily <i>store</i> information while planning or <i>controlling</i> other mental processes." (p.879).
Daneman & Carpenter (1980)	"Working memory is assumed to have <i>processing</i> as well as <i>storage</i> functions; it serves as the site for executing processes and for storing the products of these processes " (p.450).
Dixon, LeFevre, & Twilley (1988)	"... the mental system responsible for <i>holding</i> and <i>manipulating</i> information during a variety of cognitive tasks" (p.465).
Dobbs & Rule (1989)	"... the more active aspect of working memory [which is] responsible for the <i>selection, scheduling, and co-ordinating</i> of processing..." (p.500).
Foos (1988)	"... a limited capacity system that temporarily <i>stores</i> information during <i>processing</i> " (p.269).
Gick, Craik, & Morris (1988)	"... to manipulate information held in <i>short-term memory</i> while carrying out further <i>operations</i> on the stored items" (p.353).
Masson & Miller (1983)	"... responsible for both <i>holding</i> information for a limited time and for carrying out various processing <i>operations</i> " (p.314).
Morris, et al. (1990)	"Working memory tasks are those in which the person must <i>hold</i> a small amount of material in mind for a short time while simultaneously carrying out further cognitive <i>operations</i> , either on the material held or on other incoming material." (p.67).
Morris, Gick, & Craik (1988).	"... tasks in which subjects must divide their attention between ongoing <i>processing</i> and short-term <i>storage</i> " (p.362).
Salthouse, et al. (1989)	"... involves simultaneous <i>storage</i> and <i>processing</i> of information" (p.508).
Salthouse, Babcock, & Shaw (1991)	"Working memory is distinguished from... short-term memory by an emphasis on the simultaneous <i>storage</i> and <i>processing</i> of information in working memory... " (p.118).
Salthouse & Babcock (1991)	"...the <i>preservation of information</i> while simultaneously <i>processing</i> the same or other information" (p.763).

Definition of Working Memory

It appears almost obligatory in North American literature on working memory to do two things. First, one must quote Baddeley and Hitch (1974). Second, one must define working memory as involving 'temporary storage and processing' of information (Table 2), that is, to define working memory as consisting of some operations plus storage. In some instances, the concept of working memory has become so much of an assumed entity that some researchers do not even provide a definition (e.g., Engle, Nations, & Cantor, 1990; Salthouse, 1991b). The following section reviews two related operational definitions of working memory using an operation plus span framework: the Reading Span task (Daneman & Carpenter, 1980) and the complex-span definitions of Salthouse and colleagues.

The Reading Span task of Daneman and Carpenter (1980)

Of all of the quantitative/process definitions of working memory, the best known is the Reading Span task of Daneman and Carpenter (1980). The Reading Span task has been used for a long time, has been used in multiple studies, and has produced psychometrically sound data.

The original Reading Span task of Daneman and Carpenter (1980) was based on the assumption that a task using both storage and a concurrent manipulation of information is more likely to relate to other cognitive tasks, where storage measures alone do not. The Reading Span task involved presenting a sentence on a card which the person then read aloud at their own pace. The person was presented a series of these cards (from 2 to 7) and at the end of each series was asked to recall as many of the last words from each sentence as they could. Reading span was calculated as the point at which the person was able to recall only two out of three of a series length. Reading span thus defined working memory as a score on a specific task hypothesized to represent the underlying construct of working memory.


Given that the Reading Span task was designed to provide a measure of an underlying resource (a common resource hypothesis; Klapp, et al., 1983) and the predictive relationship of that resource to other cognitive operations, the efficacy of the Reading Span task must be assessed against this standard. Research using the Reading Span task has shown that it can predict reading comprehension, verbal fluency, and inferential ability (Daneman, 1991; Daneman & Carpenter, 1980; Dixon, Le Fevre, et al., 1988; Engle, Carullo, & Collins, 1991; Hartley, 1986; Masson & Miller, 1983; but not Light & Anderson, 1985). Furthermore, scores on the Reading Span task and variants of it have been shown to change with chronological age, even among school age children (Engle, et al., 1991; Gick, et al., 1988; Morris, et al., 1988; Light & Anderson, 1985). Thus, the Reading Span task does appear to provide a predictive indicator of other resource-based cognition.

However, theoretical development of how a quantitative/process model of working memory operates has not gone beyond specifying that it has processing and storage components (e.g., Just & Carpenter, 1992). There have been few direct tests of the distinction between processing and storage. One research group who have attempted to empirically separate processing from storage effects has been Salthouse and colleagues.

A generic definition of an operation plus span working memory

Salthouse and colleagues have used a complex-span definition of working memory as an independent predictor of age-related cognitive decline. However, unlike Daneman and Carpenter (1980), Salthouse and his colleagues have not restricted themselves to a verbal complex-span task. Another difference is that Salthouse and colleagues have also required that the person do more than simply read the sentence. That is, in later versions of complex-span tasks persons were required to do something with the concurrent operation (c.f., Daneman & Carpenter, 1980). This may involve, for example, verifying a

sentence as correct or calculating an equation. Also, the research by Salthouse and colleagues has not simply assumed that the operation plus storage distinction is valid; instead, this distinction has been explicitly examined. The general logic is as follows: If working memory involves two components, then measures of working memory ought to involve a measure of each component. Furthermore, separate measures of the operation and storage components might also be expected to be differentially predictive of age-associated cognitive change.

 In a study designed to examine both the operation and storage components of working memory, Salthouse, et al. (1989) proposed that working memory consisted of a storage process sensitive to the number of items presented (storage capacity) and a second process sensitive to the number of operations performed upon those items (operational capacity). They found that young adults of high ability had higher operational capacities than older high-ability adults but were no different in their storage capacity. The difference between the operational capacities of low-ability adults of different ages was much less and no differences in storage capacity was evident. From these data, Salthouse, et al. concluded that it is operational capacity, not storage capacity, that declined with age (c.f., Morris, et al., 1990).

In another similar study, Salthouse and Babcock (1991) examined the storage-operation distinction of working memory further. Salthouse and Babcock used both linear regression and path analysis methods to determine the strengths of the component relationships. Using verbal and arithmetic tasks, presented both visually and auditorily, multiple measures of working memory's hypothesized components were made. Subsequent analysis showed that, as hypothesized by other researchers (e.g., Morris, et al., 1990), both storage capacity and operational capacity explained a significant amount of variance on measures of working memory. Operational capacity and storage capacity accounted for 99% of age-associated variance on the working memory measures. When

operational capacity was statistically controlled, age-associated variance on working memory measures dropped to 1%. However, when storage capacity measures were controlled, age-associated variance on the working memory measures only dropped to 12%, indicating from the asymmetry that storage capacity mediates processing efficiency but that operational capacity directly influences both overall working memory and storage capacity. A second study replicated this result, but additionally demonstrated that simple comparison speed (of verbal and pictorial stimuli) mediated the effect of age on processing efficiency. Thus as previously concluded, it is not simply how many operations were to be performed that was predictive of working memory but "the speed with which even very elementary operations can be successfully executed" (p.775). That is, the operation component of working memory could be predicted accurately from a measure of simple comparison speed, such as comparison of line lengths (Salthouse & Babcock, Study 2).

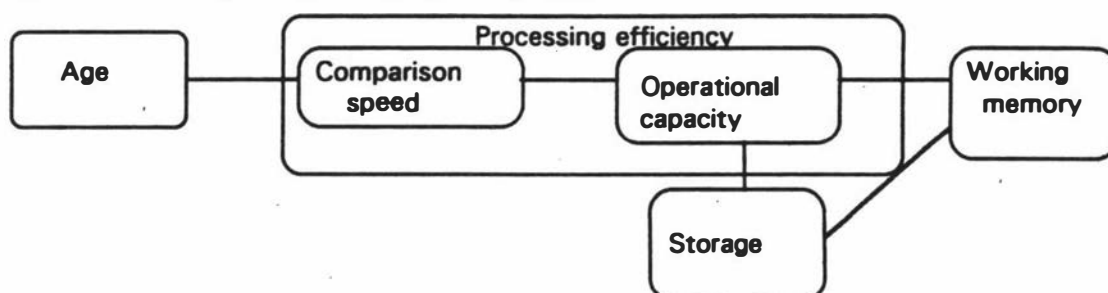


Figure 2. Salthouse's (1992) model of the relationship between storage and operations in working memory.

In a follow-up study, Salthouse (1992) replicated the path structure described by Salthouse and Babcock (1991) in which age predicted comparison speed, comparison speed predicted operational capacity, operational capacity predicted both storage capacity and working memory, and storage capacity predicted working memory (Figure 2).

In conclusion, the work of Salthouse and colleagues has provided empirical evidence for distinguishing between the storage and operational components of working memory. This evidence is in addition to the inferential evidence accumulated by studies in which complex-span

operationalisations are more predictive of other cognitive operations than simple-span operationalisations. This evidence will be discussed below.

Reliability of Complex-span tasks

Given that a major assumption about working memory is that it is a stable process in a person's cognitive makeup, any complex-span measure of working memory must necessarily produce the same or very similar estimate (or score) of working memory capacity over short periods of time. Working memory must display some test-retest reliability (Anastasi, 1982) on a complex-span task in order for that score to be representative of an enduring cognitive process.

Table 3. Reliability coefficients for various operationalisations of a storage plus processing models of working memory.

Study	Measure	r_{tt}	N
1. Salthouse & Mitchell (1989, p.512)	Computation span	.78	120
	Paper folding	.86	120
2. Babcock & Salthouse (1990, p.423).	Verbal processing	.67	80
	Spatial/processing	.61	80
3. Salthouse, et al. (1991, p.422).	Spatial	.84	50
	Numeric	.87	50
4. Salthouse & Babcock (1991, p.768).	Computation span	.90	227
	Listening span	.86	227
5. Salthouse & Babcock (1991, p.771).	Computation span	.84	233
	Listening Span	.86	233

Of the studies using the Daneman and Carpenter (1980) Reading Span task (or variants), there are very few that report any test-retest reliability figures for the task. One exception was Tirre and Peña (1992) who reported a test-retest reliability of .73 for word recall using a Reading Span task. Where test-retest reliability figures have been reported for other types of complex-span tasks, they range from .61 to .90 (Table 3). In conclusion, although there is only a bare minimum of reliability data for any operation plus span task, what data exist support a conclusion that complex-span tasks appear reliable.

However, it is important that complex-span measures of working memory (such as those used in the present research)

must demonstrate some reliability over time. Only when reliability has been confirmed can inferences be made about what a score on that complex-span task represents in terms of a stable cognitive process.

Validity Evidence for the Reading Span task.

Despite the paucity of reliability data for the use of complex-span tasks, there has been some validity evidence gathered. Validity evidence allows researchers to examine the meaning of a score on a measure in relation to a theory about the construct that the score is assessing (Anastasi, 1982).

Convergent and discriminant Validity

There have been a few studies in which the Reading Span task (or a variant of it) has been correlated against other measures of working memory and against span measures of storage alone. These studies contribute data enabling inferences about the convergent and discriminant validity of the Reading Span task. From the data in Table 4, a measure of Reading Span has been found to correlate moderately with verbally-based complex-span measures of working memory. In contrast, the Reading Span task has been found to correlate less well with simple-span measures.

Table 4. Reported correlations between the Daneman and Carpenter (1980) Reading Span task (or a variant) and other measures of working memory or of memory span alone.

Study	Task	1.	2.	3.	N
Daneman & Carpenter (1980).	1. Oral Reading Span	X			20
	2. Listening Span	.80**	X		
	3. Silent Reading Span	.88**	.75**	X	
Masson & Miller (1983)	1. Reading Span	X			29
	2. Letter Span	.07	X		
Light & Anderson (1985)	1. Reading Span	X			45
	2. Digits Backward	.34**	X		
	3. Digits forward	.26*	.51***	X	
	4. Word Span	.45***	.44***	.40**	
Hartley (1986)	1. Reading span	X			72
	2. Word naming latency	-.42**	X		
	3. Semantic verification	-.41**	.67**	X	
	4. Text recall	.03	-.28*	-.15	

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

This conclusion was also supported by studies using non-verbal complex-span measures of working memory (e.g., Babcock & Salthouse, 1990; Salthouse, et al., 1989; Salthouse, et al., 1991; Salthouse & Babcock, 1991) in which simple-span measures consistently failed to correlate as well with complex-span measures of working memory as other complex-span measures did. Consequently, it can be concluded that the Reading Span task converges with other measures of working memory.

The predictive validity of complex-span definitions of working memory for other cognitive tasks


In the previous section, evidence was presented that demonstrates that various operationalisations of working memory appear to be measuring the same construct. The present section examines the extent that a complex-span measure can predict performance on some other measure, predictive validity. Despite the incomplete definition of working memory by researchers using a quantitative/process model of working memory, there have been some interesting results produced by these operationalisations of working memory in relation to cognitive performance on various tasks.


First, the ability of the Reading Span task to predict reading processes is examined. Next, the ability of more diverse complex-span tasks to predict cognitive changes associated with chronological aging is examined. The conclusion from both subsections is that complex-span tasks demonstrate a high level of predictive validity.

The Reading Span task and its ability to predict reading processes

Daneman and Carpenter (1980) began with a question of why, if the STS is important for reading comprehension, do measures of the STS not predict performance on measures of reading comprehension. Their solution was that a new measure of the STS was needed, a measure of both storage and processing aspects. They called this measure a Reading Span task (see Just & Carpenter, 1992; MacDonald, Just, & Carpenter, 1992;

Waldrop, 1987 for a discussion of how they define working memory).

In the original Reading Span task participants were asked by Daneman and Carpenter (1980) to read a sentence aloud, remember the final word, then read the next sentence and so on, as described above. Typically persons had a mean span of 3.15 words ($SD = .93$) in this task. In their second experiment, participants were asked to verify whether sentences were true or false and were presented sentences auditorily, visually, and visually but to be read aloud. In all cases reading span was calculated as the highest number of words that were recalled for two out of three trials. Results for the three versions showed that mean spans were 2.95 ($SD = .72$), 2.38 ($SD = .70$), and 2.76 ($SD = .80$) respectively. Daneman and Carpenter found that the Reading Span task correlated well with reading comprehension measures, and from this correlation they asserted that *"If good readers use less processing capacity [of working memory] in comprehending the sentences, they should be able to produce more final words than poor readers."* (p.452). That is, using the Reading Span task, Daneman and Carpenter demonstrated an empirical relationship between measures of comprehension and working memory. 

In a replication of the Daneman and Carpenter (1980) study, Masson and Miller (1983) administered a measure of a person's ability to draw inferences in addition to the Reading Span task. It was found that working memory, assessed using a modified version of the Reading Span task, was correlated significantly with inferential ability, whereas letter span (a simple-span task) was not. The discrepancy in the correlations between the simple and complex-span measures was inferred to be due to the operation component. Thus, the storage component of working memory was suggested as being less important in predicting inferential ability and comprehension than the operation component. 

Dixon, Le Fevre, et al. (1988) extended theoretical understanding of working memory further by examining how word

span compared to the Reading Span task. They found that digit span or word span alone (STS or storage aspects) were indeed significant components of reading span. However, like Masson and Miller (1983), they found that these measures of storage were not significantly related to either comprehension or reading rate measures. However, reading span, a measure of storage and operational capacity, was significantly related to both comprehension and inferencing scores, replicating previous research on the Reading Span task and its relationship to working memory (Daneman & Carpenter, 1980; Masson & Miller, 1983).

Not all research using the Reading Span task has found that it is a better predictor of other cognitive processes than a simple-span task. Light and Anderson (1985) examined the extent of changes in both discourse and working memory and how these changes were related to chronological age. While performance on measures of working memory were similar to the levels Daneman and Carpenter (1980) found, reading span on the sentence task was not a better predictor of memory for discourse (paragraphs) than measures of digit span or word span.

In conclusion, research using the Reading Span task as a measure of working memory has typically, but not always, found a relationship between complex-span measures of working memory and of other text-based cognitive processes. That is, the use of the Reading Span task demonstrates predictive validity of that task to other cognitive tasks (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Daneman & Carpenter, 1980, 1983; Daneman & Green, 1986; Dixon, Le Fevre, et al., 1988; Masson & Miller, 1983). However, the Reading Span task and the tasks it has been shown to be predictive of are all verbally based. If the Reading Span task assesses a more general working memory, then non-verbal complex-span measures of working memory also ought to be predictive of other cognitive processes.

The predictive validity of complex-span tasks to aging .

This sub-section examines the relationship of non-verbal complex-span measures of working memory to chronological age. It has been shown in many studies (see Kausler, 1991) that with age, performance on a wide range of cognitive tasks declines. One inference drawn from this research has been that some general resource necessary to performing this wide range of cognitive tasks declines with age. An obvious choice for this general resource has been working memory (e.g., Salthouse, 1985). This subsection presents a selective review of some research using diverse definitions of working memory that has found complex-span measures of working memory to be predictive of other age related cognitive changes.

In an arithmetically-based working memory task, Campbell and Charness (1990) required participants to perform a seven step mental squaring procedure of two digit numbers. Young, middle-aged, and older adults (mean ages 24, 41, & 67, respectively) showed similar decreases in calculation errors over 200 trials. However, memory errors where components were lost or steps were missed were always greater for older than for middle-aged adults, and greater for middle-aged than for younger adults. Analysis of the proportion of errors showed that older adults increased errors in later sessions. Two explanations of the data are possible. First, there may have been reduced capacity of memory for sub-goals in older adults. Second, older adults always took longer to calculate each stage of the task. Because calculation took longer, it may be that information was lost from a STS. In either instance, the complex-span measure of working memory was clearly related to age.

In a further study using an arithmetically-based measure of working memory, Dobbs and Rule (1989) examined persons' ability to change what was stored in working memory by assessing recall of auditorily presented digits at lag 0, 1, or 2. Dobbs and Rule used the mean number of digits recalled to the first error in recall as the indicator of working memory, a measure which showed differential age related

declines as the lag was increased. Thus, again a complex-span measure which examined both storage and operational components was related to chronological age.

4 Salthouse, et al. (1989) predicted that working memory represents a general resource which declines with age. From this prediction, it is clear that measures of working memory ought to predict age-related declines in other cognitive tasks better than chronological age alone. In a test of this hypothesis, Salthouse, et al. used a computation span measure of working memory to compare to reasoning and paper folding tasks. The correlations of computation span with reasoning and paper-folding tasks were .48 and .38, respectively, both significant at $p < .01$. Furthermore, when variability due to the computation span task was removed, age-related variability on measures of reasoning dropped from .28 to .12 and on paper-folding from .28 to .16. This drop in variability indicates that the computation span measure of working memory is an independent predictor of other cognitive tasks.

Finally, Salthouse and Skovronek (1992) presented young and older subjects with a cube-comparison task designed to assess the availability of intermediary information at various stages of processing. Young adults again scored higher on three measures of working memory and on the other concurrent cognitive tasks than older adults.

D In addition to verbally-based operational definitions of working memory, using non-verbal definitions of working memory has shown that working memory scores remain predictive of other cognitive operations (see also, Engle, et al., 1990; La Pointe & Engle, 1990; Turner & Engle, 1989). As Salthouse (1990) states, it appears to matter little how working memory is assessed (sentence span or computation span) nor whether presentation is auditory or visual. Overall, the non-verbal complex-span tasks have demonstrated what the Reading Span tasks have: that it is not simply storage or operational capacity singly that measures working memory, rather a

combination of both storage and operational capacity appear to provide the best assessment of working memory.

Conclusions

This section began with an examination of definitions of working memory. It has been shown that there is some utility in the operation plus storage distinction. However, use of this distinction also requires that attempts be made to assess these dimensions separately to avoid the assumption that both operation and storage are equipotent in their impact on other cognitive tasks. Also, there are reliability, convergent validity, and predictive validity data supporting the use of a complex-span task to assess the quantitative/process model of working memory. Finally, scores on the complex-span task are typically better predictors of other cognitive task performance than scores from a simple-span task (but not always; see Light & Anderson, 1985). However, further use of a complex-span task must explicitly check the reliability of the measure, especially as complex-span measures are theorized as representing a relatively stable cognitive resource.

Overall, the evidence reviewed thus far allows two conclusions to be drawn. First, the operation plus storage distinction is a useful one in assessing working memory. Second, operationalisations of this operation plus storage distinction are robust across different modalities.

A THEORETICAL FRAMEWORK FOR THE QUANTITATIVE/PROCESS MODEL OF WORKING MEMORY

Having established the utility of the complex-span task as an operational definition of the quantitative/process model of working memory, attention can now be turned toward examining what this operational definition contributes to the theoretical understanding of the quantitative/process model of working memory. This section examines three aspects of working memory: preservation of content information;

preservation of order information; and, the role of temporal factors in the operation of working memory.

Preservation of Content in Working Memory

It almost seems redundant to write that working memory preserves intermediary and end products of processing. Since the time of Ebbinghaus and others (Murray, 1980), memory of all types has been assessed by what is preserved and how well what is preserved can be recalled. For example, a typical answer to the question of how someone's memory is performing usually involves some comment about how *much* is remembered now, compared to some previous time. Memory and preservation of content seem synonymous.

Yet there is an accumulating body of evidence that suggests that content does not totally explain the role of working memory in other cognitive tasks (e.g., Salthouse, et al., 1989). This evidence comes from two sources: from comparisons of working memory scores obtained on complex-span versus simple-span tasks; and from path analysis of the relationships between measures of working memory components.

As has been reviewed so far, content-only information using a simple-span task does not usually predict other cognitive operations as well as content information obtained while performing a complex-span task (e.g., Daneman & Carpenter, 1980; Dixon, Le Fevre, et al., 1988; Masson & Miller, 1983; c.f., Engle, Cantor, & Carullo, 1992). This suggests that the variability introduced into the content measure of a complex-span task by the concurrent operation adds predictive power to the score beyond that given by a storage-only score. That is, when working memory capacity is measured using operation plus storage, the resultant score is a better predictor of other cognition than a score of capacity based on storage alone. However, a capacity score on an operation plus storage task is clearly an amalgam of operational plus storage capacity. To test the impact of a concurrent operation on working memory, there is a clear need to obtain independent estimates of both storage and operational capacity.

Salthouse and Babcock (1991) used independent estimates of storage and operational capacity to examine how well they each correlated with an overall measure of working memory. In their first and second experiments, Salthouse and Babcock obtained correlations ranging from .39 to .56 between measures of storage and operational capacity (corresponding to how many operations can be performed in a given time). In predicting age-related differences in other cognitive tasks however, Salthouse and Babcock found that content measures were not as effective as operational capacity measures in predicting age-related variance. These results reinforce that, although what is stored has been traditionally perceived as of greatest importance, that it is also worthwhile to examine operational capacity (see also Salthouse, et al., 1991; Salthouse & Skovronek, 1992).

Order Information in Working Memory

A key feature of operational capacity is being able to maintain the order of what is recalled. Order information appears to be co-ordinated in working memory in two ways in a complex-span task. First, there is the co-ordination of the order of the intermediary products of a concurrent operation. For example, in the calculation task of Campbell and Charness (1990), persons had to maintain up to seven numbers in order to achieve the final result. However, not only did the content of the products have to be maintained, but the order that those intermediary products were in must also be preserved in order to obtain a correct final result. Clearly, in performing a processing task, there must be some method of 'tagging' the order of the intermediary products of the processing task.

Second, as was discussed in Chapter 2, the order of items in the storage component may also need to be maintained by working memory. For example, Babcock and Salthouse (1990) required the numbers from the secondary (storage) task to be recalled in order (see also Brainerd & Kingma, 1985; Daneman & Carpenter, 1980, 1983; Salthouse, 1991b, 1992; Salthouse &

Babcock, 1991; Salthouse, et al., 1989; Salthouse & Skovronek, 1992). Thus, in both complex and simple-span tasks, being able to preserve the order of items is a necessary component to the recall task.

Conclusion

In conclusion, the simple and complex-span tasks both clearly require some ability at maintaining order information. Whether the sequencing is of intermediary products necessary for performing an operation successfully, or of items necessary to recalling a span correctly, working memory is clearly very involved in the preservation of order information.

Temporal Factors in Working Memory

From the complex-span task, it is also clear that the quantitative/process model of working memory is also temporally limited. This subsection explores the nature of those temporal limitations.

Complex-span tasks, like simple-span tasks, are typically presented as rapidly as the information can be assimilated, and then recall is also usually as rapid as can be managed without being incorrect (e.g., the Reading Span task of Daneman & Carpenter, 1980). The implication is that working memory is temporally limited, that is, only maintains items for a brief period of time.

With complex-span tasks, it is also difficult to control stimulus presentation because it obviously takes some people more time than others to complete an operation. Using the Reading Span task (Daneman & Carpenter, 1980) as an example, it may take person A 800 ms to verify each sentence correctly. It may take person B 1200 ms to verify the same sentence. Thus, if presentation time were only for 1000 ms person A, but not person B, would be able to verify the sentence. At the end of several verifications, person A may have a span of 3 items. Person B may have a span of 4 items. From this example it is unclear whether person A's span was

smaller than person B because of some smaller working memory or because person B simply did not bother to do the additional verification task. There is thus an obvious confounding of the verification condition with the recall condition. The simplest way to avoid this is to allow each person to take as long as necessary to verify each sentence, as is done with most variants of the Reading Span task. However, in allowing verification times to be self-paced, there is an onus on the researcher to also measure any verification time differences, something which is not typically done (e.g., Daneman & Carpenter, 1980; Salthouse, et al., 1989). Measurement of the time it takes persons to verify each sentence is thus a necessary variable to consider in the present dissertation.

However, in measuring the time to perform a complex-span task, one must be precise about what exactly is being measured. That is, what information does verification time provide? Borrowing the reasoning used by Salthouse (1988a), this verification time may represent the total time to successfully complete a set of operations. If the same cognitive operation takes different times for different persons, this represents different processing rates of information in working memory. Second, different verification times may reflect differences in sequencing the operations occurring at the time. That is, there may be a differential ability to preserve the place one is up to in a sequence of operations; some subjects may need to engage in backtracking, and ultimately require longer to complete the task. Finally, different verification times may represent differences in some working memory capacity that is necessary for performing the complex-span task. This hypothesis implies a finite amount of storage capacity (Salthouse, 1988a) that can be allocated for the working memory process.

In a sense, each of these three explanations explain the meaning of verification time in terms of processing rate, order manipulation, and content of some component process of working memory. Given that working memory itself is being

explained in these terms, such a second-order explanation of the temporal component is extremely recursive.

One possible solution to this problem is to ask the question of what the verification time of a sentence means in and of itself. That is, before assuming that the time to verify a sentence represents some component of working memory, it is necessary to first empirically establish a relationship between verification time to perform a complex-span task and the other aspects (content and order) of working memory. If this verification time is not found to be related to other working memory processes, then there is little value in trying to incorporate a measure of a more general processing time into a working memory model.

Conclusions

The present section has examined what theoretical variables the empirical complex-span task implies for the theory-based quantitative/process model of working memory. Consideration of the complex-span task suggests that working memory involves preservation of content information; working memory involves maintenance of order information, whether within an operation or between the span items; and working memory involves some temporal aspects -- *processing time*. The conclusion of this section then is that aspects of content, order, and processing time may be critical in assessing the operation of the quantitative/process model of working memory.

SUMMARY

This first section of the present chapter reviewed evidence for the reliability and validity of complex-span task operationalisations of working memory. The evidence favours an interpretation that the complex-span task is a better method of assessing working memory than a simple-span task alone. However, this conclusion is not without some caution. Specifically, the basis for this conclusion has been that complex-span measures of working memory correlate better with

other tasks than simple-span measures do. However, there has been little direct empirical comparison between complex and simple-span tasks to support an inference that, in and of itself, a complex-span task is necessarily a better measure of working memory per se.

Also, despite the operational success of the quantitative/process definition of working memory, there has been little specification of what either an operation or span consists of since the original Baddeley and Hitch (1974) definition of working memory. The second section of the present chapter identified at least three key variables implicated in the operation of the quantitative/process model of working memory: content, order, and processing time. Of these three, processing time is potentially the most troublesome. Rather than trying to define processing time in terms of what is occurring, the present research will first explicitly test whether the inferred relationship between processing time and the other two variables can be empirically demonstrated. While this may seem to be postponing discussion of the nature of processing time, it will be shown later that this choice is empirically justified.

CHAPTER 4: SUBJECTIVE ON-LINE META-MEMORY MEASURES

A third dimension of working memory is defined by the distinction between measuring working memory by performance (objective) or by self-report (subjective) methods. This distinction has been captured in the past by referring to performance and meta-memory measures respectively (e.g., Cavanaugh, 1989; Dixon, 1989; Kausler, 1991). However, while it is tempting to refer to overt measures as *what we know* and to meta-memory as *what we know about what we know* (e.g., Brown, 1975), this distinction may not be as easily made empirically. Chapter 4 specifically discusses the measurement of meta-memory and the relationship between meta-memory and working memory.

There have been many research articles and reviews published examining meta-memory over a diverse range of perspectives. Meta-memory has been examined both theoretically (e.g., Cavanaugh, 1989; Cavanaugh & Perlmutter, 1982; Czerniawska, 1983; Dixon, 1989; Flavell, 1971; Flavell & Wellman, 1977) and in relation to memory performance (e.g., Barclay, 1981; Cavanaugh & Borkowski, 1979; Cornolodi, Gobba, & Mazzoni, 1991; Hasselhorn & Hager, 1989). There have also been many methods of operationalising and assessing meta-memory, methods ranging from questionnaire assessment to online pre-estimates (e.g., Dixon, Hertzog, & Hultsch, 1986; Dixon & Hultsch, 1983a, 1983b; Dixon, Hultsch, & Hertzog, 1988; Gilweski, Zelinski, & Thompson, 1983; Hertzog, Dixon, & Hultsch, 1990; Hertzog, Dixon, Schulenberg, & Hultsch, 1987; Hertzog, Hultsch, & Dixon, 1989; Hultsch, Hertzog, & Dixon, 1987; Kausler, 1991; Leonesio & Nelson, 1990; Nelson & Narens, 1990). Meta-memory has also been examined in relation to memory performance declines associated with age (e.g., Agrawal & Chabra, 1989; Baillargeon & Neault, 1989; Bruce, Coyne, & Botwinick, 1982; Cavanaugh & Murphy, 1986; Cavanaugh & Poon, 1989; Devolder, Brigham, & Pressley, 1990; Devolder & Pressley, 1989; Lachman, Lachman, & Thronesbury, 1979; Loewen, Shaw, & Craik, 1990; Murphy, Sanders, Gabrieshki, &

Schmitt, 1981; Murphy, Schmitt, Caruso, & Sanders, 1987; Perfect & Stollery, 1993). Meta-memory changes have been examined in relation to clinical disorders such as depression (Niederehe & Yoder, 1989), amnesia (Janowsky, Shimamura, & Squire, 1989; Parkin, Bell, & Leng, 1988), multiple sclerosis (Beatty & Monson, 1991), and temporal lobe epilepsy (Prevey, Delaney & Matson, 1988). Finally, meta-memory has often been examined in relation to learning strategies of children (Kontos, Swanson, & Frazer, 1984; Kurtz & Weinert, 1989; McNamara, 1980; Pressley, Borkowski, & O'Sullivan, 1984; Weed, Ray, & Day, 1990). However, in all this research, *meta-memory has never been examined in relation to working memory.*

In considering the research on meta-memory, one is also easily drawn into using such terms as awareness, knowledge, mindfulness, cognizance, attention, and other mind-related descriptors. There is a very real danger in using these terms that they can be used as literary, yet imprecise, descriptors of the construct of meta-memory. It is clearly unacceptable to define meta-memory in terms of synonyms. There needs to be a language capable of describing both the theoretical and empirical aspects of meta-memory. However, the author is not convinced that such a language exists that is simultaneously able to describe the richness of meta-memory without becoming vague and to also capture the components of meta-memory without being overly reductionistic. Any discussion of meta-memory involves at least acknowledgment that one is discussing the knowledge one has about one's knowledge -- a clearly paradoxical statement in which a member of the subset of one's knowledge (meta-memory) is described using the same language describing the set to which meta-memory belongs (mind). This conflict of logical types (Watzlawick, Weakland, & Fisch, 1974) is intuitively demonstrated by the difficulty persons have in conceptualizing what meta-memory is, and in the literature in which many synonyms are used, most of which lack conceptual substance. Any theory of meta-memory must at least describe meta-memory in language capable of avoiding confusion with metaphysical terms such as knowledge and

awareness, that is, such a language must attempt to provide external or at least independent referents (see also Cavanaugh, 1989). As will be seen below, even this minimal expectation is only partially met.

The following sections examine theoretical aspects of meta-memory, operational aspects of meta-memory, and the assumptions made in meta-memory research. However, the main goal of this review is to define parameters for considering whether persons are able to monitor their working memory. That is, *are subjective measures of working memory able to inform one about the nature and operation of working memory in the same or different ways that objective (performance) measures can?*

The Theoretical Basis of Meta-Memory

In the first use of the term meta-memory, Flavell (1971) defined it very broadly as including both what a person knows about their personal memory and what they know about remembering in general. Thus, meta-memory was defined as referring *"to an individual's knowledge of and awareness of memory, or of anything pertinent to information storage and retrieval."* (Flavell & Wellman, 1977, p.4). Meta-memory, according to Flavell and Wellman, involves personal knowledge about *item characteristics, task demands, personal attributes* and *available memory strategies*. Together, the first two of these features determine one's assessment of the difficulty of the memory task, and the latter two the ability of the person to complete the task respectively. For example, one might have to learn a series of 20 mathematical formulae. The characteristics of the items are such that they are not familiar to the learner and so are perceived as difficult. The task demand that 20 items be learned is also such that one might judge the task to be difficult. However, one might have learned lists of formulae successfully before (personal attributes) and might have found that starting with basic equations and using integral calculus to derive further formulae was a successful strategy. So, overall, when

considering task difficulty and personal ability together, one forms a meta-memory judgment about one's probable memory performance.

Tulving (1985) makes a similar distinction to the difficulty/ability distinction of Flavell and Wellman (1977) in distinguishing between *noetic* and *autonoetic* awareness. Noetic awareness is what one knows about one's knowledge base and the needs of a memory task. Autonoetic awareness is less easy to define. Autonoetic awareness involves a subjective perception of the state of one's current memory; knowing the current state of one's cognitive system.

In a similar manner to Flavell and Wellman (1977), Dixon (1989) defined meta-memory as

a term representing one's knowledge, perceptions, and beliefs about the functioning, development, and capacities of (1) one's own memory and (2) the human memory system. It includes knowledge, perceptions, and beliefs about the demand characteristics of various tasks or situations, the availability and employability of relevant strategies and aids, and other memory-relevant characteristics of the persons themselves. (p. 396)

Dixon then noted that even this definition is at times supplemented and altered to suit various researchers' particular purposes. However, Dixon's usage of meta-memory extended the Flavell and Wellman definition by applying it to other than only memory development (by implication in children) and by describing the definition of meta-memory in perceptual rather than external terms. This latter point means that the nature of meta-memory is described rather than the factors that produce meta-memory.

Memory Knowledge

Memory knowledge (Dixon, 1989) involves one's awareness of the static store of strategies that might help performance on a memory task (Cavanaugh, 1989; Flavell & Wellman, 1977). For example, when memorizing a list of facts about the Treaty of

Waitangi, one might be aware that method of loci is a strategy useful for remembering lists. Cavanaugh defines this knowledge that the method of loci, for example, is a possible strategy for remembering these facts as systemic awareness. This type of awareness is essentially knowledge based -- what Tulving (1985) termed noetic awareness. When a strategy has been selected and used, the effectiveness of that strategy can then be evaluated. The awareness (autonoetic awareness) a person forms from this evaluation of the effectiveness of a memory strategy refers to memory perceptions.

Memory Perceptions

Memory perceptions (Dixon, 1989) involve being able to evaluate the products of one's remembering against the demands of a memory task. For example, the demands of remembering ten facts about the Treaty of Waitangi can be compared against the product, or facts recalled, after using the method of loci. If one had ten facts to learn, and one recalled ten facts, then a perception that one had remembered all of the facts might be formed. If one only recalled two facts, then the perception may be that one had not remembered very much. This is what Cavanaugh (1989) described as on-line awareness of current processing and what Flavell and Wellman (1977) described as one's knowledge of probable memory performance. These perceptions of the current state of one's remembering are necessarily a component of any executive processing (Cavanaugh). Unlike memory knowledge memory perceptions are dynamic, changing as one becomes aware of one's level of retrieval. The emphasis in memory perceptions is on current awareness of the products of one's past memorizing.

Memory Beliefs

When time has elapsed between recall of those products of a memory strategy, memory perceptions themselves become a memory. So, when a person is then asked to report their memory of their past recall, one begins to assess memory beliefs (Dixon, 1989). For example, if one learned a list of ten facts two years ago, and you were asked now to report on

how effective the method of loci was, the report that you give will be based on your memory of your memory perceptions two years ago. Consequently, your report of past memory perceptions will reflect both the actual memory perception experienced in the past, and the current beliefs (or awareness of) that past remembering experience.

Epistemic Awareness

Cavanaugh (1989) defined the combined general knowledge of many past on-line 'awarenesses', or what Dixon (1989) called memory beliefs, as epistemic awareness. Epistemic awareness "...is [a general] knowledge about one's own knowledge base....Epistemic awareness occurs during the retrieval process, when the judgment about knowledge is made..." (Cavanaugh, 1989, p.419). One particular type of epistemic awareness is whether memories for past memorizing (e.g., memory for memory perceptions) are real or generated (Cavanaugh). This type of epistemic awareness, which the present author calls *source epistemic awareness*, enables one to decide if the recall of knowledge about past use of memory strategies is accurate, that is, whether it was personally experienced or not. The degree of source epistemic awareness that a person possesses determines the extent that recall of past on-line awareness is affected by present personal beliefs about memory. For example, if you are asked to recall how well you recalled a list two years ago, you may only be vaguely (if at all) aware of the list and of recalling it. But you believe that the method of loci works well, and you are aware that it has enabled you to recall many other more complex lists. So, you report that your recall of the particular list two years ago was greatly enhanced by the method of loci. However, when further questioned, you become aware that you do not really recall the level of memory perceptions you had two years ago, rather that your recall of that awareness is based more probably on your current positive beliefs about the method of loci. It is this experiencing of the source of one's knowledge that is epistemic awareness. Epistemic awareness is invaluable in

determining how well memory performance is later reported, as either a veridical report, or the extent that beliefs about remembering have coloured one's current recall of the past.

Recall that Dixon (1989) suggested three components of meta-memory; knowledge, perceptions, and beliefs. It has been described that the first two of these categories relates to systemic and on-line awareness respectively (Cavanaugh, 1989). Knowledge and perceptions are present-oriented, that is, are levels of awareness. Later recall of one's perceptions may be confounded by intervening (temporally) factors, which form beliefs that bias the accuracy of one's reporting of past perceptions. Epistemic awareness, again present-oriented, is a type of meta-memory monitoring that allows one to assess the degree that self-reports of past perceptions are veridical, or alternately are affected by beliefs.

Dixon (1989) thus defines meta-memory as a construct made up of several theoretically based smaller parts. So when one is describing or assessing meta-memory, it is crucial to specify whether it is memory knowledge, memory perceptions, memory beliefs, or source epistemic awareness that one is talking about (see also Cavanaugh, 1989; Flavell & Wellman, 1977).

While the description provided by Dixon (1989) and Cavanaugh (1989) allows the theoretical boundaries of meta-memory in the current study to be explicated, these boundaries are devoid of any other empirical support apart from the introspective data we can all generate. To more clearly define meta-memory it is important to also examine how meta-memory is operationally defined.

The Operational Basis of Meta-Memory

Kausler (1991) defines meta-memory as

knowing what one's own memory system can and cannot accomplish. Through such knowledge we are able to assess... how much new episodic information we can assimilate in a given period of time,...how likely

previously encoded information is retrievable,... what memory strategies we have available and how to deploy those strategies... (p.534).

These three facets of meta-memory are labelled off-line evaluation of memory proficiency, on-line evaluation of memory proficiency, and the monitoring of memory performances respectively.

The present dissertation was constructed to be about examining facets of working memory, one of which is that working memory has a short duration. Thus in examining one's meta-memory for a particular working memory, one is also examining, at least in the first instance, an awareness of a momentary phenomena. In Dixon's (1989) terms the corresponding momentary awareness of transient working memory performance must be a perception of working memory performance, which after a brief time period and before the next task becomes a memory belief. Empirically, because these perceptions, which then form beliefs, occur close and in response to actual performance, in the present dissertation one is operationally examining on-line meta-memory. Thus, only a review of empirical work on on-line meta-memory will be presented here.

On-line meta-memory evaluation

On-line evaluation of memory proficiency examines persons' ability to either pre- or post-dictively evaluate their performance on a memory test (Kausler, 1991). Assessment of this meta-memorial ability has typically consisted of providing the person with an example of the memory task, then asking them how many of the stimuli they will recall or recognize. For example, a person might be told that they are to be shown 20 items in a free-recall procedure and asked to estimate how many of those items they would recall after a single presentation. Their response is termed an on-line meta-memory pre-estimate. On-line meta-memory pre-estimates can also be divided into estimates for already learned material (Feeling-of knowing, FOK), or estimates for as yet

unknown material (Ease-of-learning judgments; EOL; Costermans, Lories, & Ansay, 1992; Nelson, Gerler, & Narens, 1984). Because the present dissertation is only concerned with to-be-learned material, only research examining this category will be reviewed.

Meta-memory for to-be-learned material

Using on-line pre-estimates, it has been shown that pre-estimates are generally accurate on free-recall tasks but that younger persons tend to under-estimate and older persons tend to over-estimate performance (Bruce, et al., 1982; Devolder, et al., 1990; Lovelace, 1984; Murphy, et al., 1981; Rebok & Balcerak, 1989; c.f. Perlmutter, 1978). Pre-estimates have also been shown to be accurate for intellectual performance tasks (Lachman & Jelalian, 1984).

There are two important points to notice from the available literature. First, examination of on-line meta-memory pre-estimates has usually involved comparisons of accuracy between groups rather than examination of veridicality within groups. Second, none of the pre-estimates were of working memory performance. There is thus a clear gap in the literature regarding meta-memory and working memory. Given that working memory measures are accorded some importance as an indicator of one's ability to perform other cognitive tasks, it is surprising that persons' ability to monitor their working memory has not been examined. Thus, the present dissertation, in developing a fuller description of working memory performance, considers it proper to also begin to develop a description of how working memory is monitored.

The final point to consider in using an on-line meta-memory measure of working memory relates to what theoretical aspect of meta-memory it is a measure of. Unfortunately, the theoretical development of meta-memory is not at a point where one can definitively claim exactly what aspect of meta-memory is being assessed. In respect of all the above points, any examination of meta-memory in relation to working memory must be considered exploratory.

Assumptions of Meta-Memory Research

As in any research, there are always levels at which knowledge is taken as accepted, the assumptions of the research. The major assumption made in examining self-report data has been that persons are capable of accurately reporting their cognitive state. There are two issues contained in this statement. First, self-reports must accurately reflect some internal state. There is evidence that persons are able to provide accurate data relating to their cognition and to meta-memory in particular. For example, Nelson and Narens (1990) review research using psychophysical principles which shows that meta-memory in general and 'Feeling-of-knowing' (FOK) judgments in particular can be accurately reported by persons. Also, Leonesio and Nelson (1990) reported that Ease of Learning (EOL) judgments prior to performance have been able to predict subsequent performance. Second, irrespective of whether reporting is accurate, there needs to be some demonstration that subjective measures do actually reflect some internal cognitive state or process. Again, there is some evidence that suggests that it can be safely assumed that persons are capable of examining their own perceptual (and cognitive) processes when placed in Signal Detection paradigms (e.g., Pastore & Scheirer, 1974).

It is also necessary to assume that self-reporting of subjective states be considered data in the sense that other objective measurements are (Ericsson & Simon, 1980). Also, even though correspondence between performance and verbal reports may be low, this in itself can inform the researcher about the state of persons' ability to report on their cognition (e.g., Gilewski, & Zelinski, 1986; Herrmann, 1982; Herrmann, Grubs, Sigmundi, & Grueneich, 1986). As data show in the area of aging research, where meta-memory estimates are less than veridical, they are still of use in untangling hypotheses by comparing them across groups for qualitative and quantitative changes (e.g., Devolder, et al., 1990; Lovelace, 1984).

Finally, in examining the nature of meta-memory one is also necessarily examining a bi-directional relationship between performance and monitoring. Nelson and Narens (1990) describes this bi-directional relationship as between control and monitoring processes. That is, meta-memory affects memory and memory affects meta-memory. Indeed some research suggests that with the onset of Alzheimer's Dementia, there is a lower awareness of memory performance; that is, meta-memory declines in conjunction with other memory (McGlyn & Kasniak, 1991). Thus, as noted by Dixon (1989), there is an assumption that there exists a bi-directional relationship between meta-memory and memory, a point examined in further detail in Chapter 5 below.

Conclusions

From the above section examining theoretical aspects of meta-memory, it was concluded that meta-memory involves memory knowledge, perceptions, beliefs and source epistemic awareness. The second section examined operationalisation of on-line memory monitoring measures. The third section of the present chapter examined the assumptions of meta-memory research that persons are able to accurately report on their cognition, that such reporting is useful, and that there exists a bi-directional relationship between meta-memory and memory.

The question asked at the beginning of the chapter was whether persons could monitor their working memory. When the theory and research bearing on this issue is considered, there is a clear need for some empirical investigation of the ability of persons to specifically *predict the operation of working memory*, and perhaps even of its components. However, given the absence of prior research, the present research into the relationship between working memory and meta-memory is fundamentally exploratory with only a few guiding principles.

CHAPTER 5: A RATIONALE FOR INVESTIGATING THE TWO MODELS OF VERBAL WORKING MEMORY

At the end of Chapter 4, three dimensions or ways of measuring working memory were presented. It has been shown that working memory can be described as a set of structures or as a set of processes; these descriptions are referred to as the qualitative/structural and quantitative/process models, respectively. Also, it has been advanced that working memory might be able to be described at the level of meta-memory. However, at this point, it is apparent that very little overlap at either operational or theoretical levels exists between these dimensions. This lack of overlap begs two questions in relation to verbal information processing: *Are the two models of working memory in fact examining the same construct?* and *Can persons monitor their working memory?* Chapter 5 presents the rationale for examining the relationships between the two models of working memory, and between working memory and meta-memory.

In Chapters 2 and 3 of the present dissertation, two models of working memory were presented. The qualitative/structural model of working memory was discussed in terms of its components and its operational defining features. Specifically, the Phonological Loop was shown to be represented by the phonological similarity and word-length effects and in relation to the impact of articulatory suppression. The quantitative/process model of working memory was defined in Chapter 3 as being composed of operational and storage components. From the quantitative/process model of working memory, it was concluded that the dimensions of content, order, and processing time are important aspects of working memory. Finally, research with both the qualitative/structural and quantitative/process models of working memory has found that the scores obtained on measures of working memory are predictive of other 'higher' cognitive processes.

Before outlining the relationships between the two models of working memory, a fundamental question that must be addressed is why the quantitative/process and qualitative/structural models can be described as two models and not simply as two components of the same model. The present dissertation argues that there is sufficient scientific strain to require resolution of the relationship between the models and that there is sufficient evidence to consider these as separate models of the same construct. The distinctiveness can be demonstrated by the history, literature, application, simulation and operational definitions of the models.

Historically, both models originated in the work of Baddeley and Hitch (1974) who proposed that working memory involves storage and processing. However, the subsequent development of working memory is acknowledged by some researchers as being along similar but increasingly diverging lines of inquiry (e.g., Gathercole, 1994; Hasher & Zacks, 1988). Indeed, Baddeley and Hitch (1994) propose that there are three separate uses of the term working memory at present: the two uses described herein plus computational models (e.g., Newell, 1990, cited in Baddeley & Hitch, 1994).

The literature also reflects that these two developments of working memory can be considered two models. For example, Just and Carpenter (1992) comment that *"The working memory in our theory corresponds approximately to the part of the central executive in Baddeley's theory...The working memory in our theory does not include modality specific buffers"* (p. 187) These statements clearly note that the authors consider there to be two theories of working memory and that, while they acknowledge a relationship between the two, they reject the concept of subsystems, a concept integral to Baddeley's (1986) model. Baddeley (1992c) also comments that *"While Carpenter and her colleagues would not adopt the particular model of working memory just described, they do not deny the existence of more peripheral systems such as the phonological loop"* (p. 287) Again, here is a clear statement that the models are distinct. In addition to literature implying

distinctiveness at the levels of theory and models, La Pointe and Engle (1990) also comment that " *...the simple word span and the reading span are inherently different...*" (p. 1119) with the implication that task differences reflect some inherent difference between models. Overall then, the literature appears to draw distinctions between these two 'models' at both theoretical and task levels.

When the 'real world' applications of these two models are examined, there are again clear differences between them. The quantitative/process model is typically applied to the measurement of individual differences using psychometric principles which require establishing links between constructs such as working memory and reading comprehension (e.g., Daneman & Carpenter, 1980). In contrast, the qualitative/structural model typically is applied in single case designs in neuropsychological studies designed to fractionate components of working memory within an individual (e.g., Baddeley, et al., 1987).

It is also clear from the simulations of these two models that there are some particular differences. Burgess and Hitch (1992) simulated the qualitative/structural model's ability to explain the phonological similarity, word length, and articulatory suppression effects. Their model is not able to deal with sentential input nor to predict cognitive processes other than working memory. In contrast, Just and Carpenter (1992) present a model that is able to relate a working memory capacity to sentence comprehension and that claims to simulate sentential working memory capacity.

Finally, how the two models handle verbal information are prototypically defined by two separate tasks. The model explicated in Baddeley (1986) has typically been operationalised using a single stimulus item, usually a letter or word, when defining verbal working memory. A set of these letters or words are typically presented briefly to the participant who then rapidly recalls them in their correct order. In contrast, the quantitative/process model has had no consistent operational definition. The only operational

certainty is that the task will have an item to recall plus a task to perform concurrently. It is important to reiterate here that these tasks are not simply two ways of measuring the same construct, but are more usually perceived as inherently different (e.g., La Pointe & Engle, 1990).

In sum, an alternative viewpoint might be that what have been termed two models here are in fact but two facets of the same model. This perspective, in the opinion of the author, is difficult to substantiate however when considering the differences between the two as evidenced by their history, the literature about them, how they are applied, how they are simulated, and how they are operationally defined. It is not to be taken, however, that the present work claims the two models are irreconcilably distinct either. Rather, the present work seeks to clarify the relationship between the two models with respect to how they each account for storage of verbal information. It is also important to return to a point made in the summary of Chapter 2 that the simple task used here in no manner purports to be a complete representation of the qualitative/structural model. Thus, one cannot make conclusions about the qualitative/structural model *in toto* based on this task. In contrast, the complex span task does appear to completely define the quantitative/process model of working memory, thus illustrating a further difficulty in examining how these models handle verbal information: they are in some respects different in scope.

To begin the process of clarification, the premise adopted is that these are two separate models, a premise based on the available, although at times unclear, evidence. It is thus a scientifically sensible task to investigate how these two models account for verbal information in order to begin the larger examination of how the complete qualitative/structural model relates to the quantitative/process model or approach.

Relationships Between the Qualitative/Structural and Quantitative/Process Models of Working Memory

Theoretically, some consideration has been given as to the links between the qualitative/structural and quantitative/process models of working memory. Just and Carpenter (1992) have recently acknowledged that working memory must be more clearly defined and that the definition must in some way relate to Baddeley's model. In considering the relationship between the two models, Baddeley (1992b) described them as complementary with each having strengths and weaknesses such that

...the psychometric [quantitative/process] correlational approach has the advantage that it can tackle...the Central Executive, and can furthermore work directly on problems of practical significance, such as reading comprehension or the reasoning tasks used in tests of intelligence. The weakness of this [quantitative/process] approach lies in its reliance on complex working memory tasks that have a somewhat arbitrary construction and do not readily lend themselves to a more detailed analysis of the component processes. (pp. 556-557)

That is, it has been acknowledged that the two models of working memory ought to be related and probably are somehow. In a clear statement of what the relationship between the two models might be, Just and Carpenter claim that their research investigates the Central Executive, and that the Phonological Loop is not a part of the process that they investigate, a claim further promulgated by Gathercole and Baddeley (1993) and Hasher and Zacks (1988).

Yet, when one considers the nature of the tasks used in both models, it is apparent that the tasks used by Just and Carpenter are visually presented reading tasks that require concurrent memorization of words (e.g., Daneman & Carpenter, 1980), the same type of tasks that are used to define the Phonological Loop. And so, while there is some theoretical

consensus about what the relationship between the qualitative/structural and quantitative/process models of working memory is, the basis for this evidence is uncertain. Speculatively, it appears convenient to suggest that because the Central Executive is said to be an area of residual ignorance (Baddeley, 1986) that this is the domain of the quantitative/process model of working memory. It seems to the author that it might be just as sensible to speculate that the Central Executive is concerned with the operation aspect of the quantitative/process model and that the Phonological Loop is the same as the storage area of the quantitative/process model. This would imply that both 'models' account for storage of visually presented verbal information in the same way using near identical processes. This more parsimonious explanation avoids accepting the presumption that one model uses the Central Executive primarily to store verbal information and the other uses the Phonological Loop. That is, one has the option of describing the two models as simply two separate components (the received wisdom in the literature) or of demonstrating that the two models in fact both use the Phonological Loop in a similar way. So, in summary it would seem that there is a presumption that the quantitative/process model of working memory is involved with the Central Executive. However, the empirical basis of this presumption is uncertain.

Conversely, there has been little cognizance in the literature that there is a need to describe the storage of verbal information in the qualitative/structural model in terms of the processes operating in quantitative/process models or approaches to working memory. This omission is understandable if one accepts the earlier assertion that the quantitative/process model is subsumed in the Central Executive (A.D. Baddeley, personal communication, April 19, 1993; Gathercole & Baddeley, 1993; Hasher & Zacks, 1988; Just & Carpenter, 1992). However, given that the review in Chapter 2 highlighted that order information is important to the qualitative/structural model, and that order has been

suggested as a dimension of the quantitative/process model, again logic dictates that the relationship between how the two models store visually presented verbal information be examined more carefully from both perspectives.

In terms of theory, there appears to have been little serious attempt to define the relationships between how the qualitative/structural and quantitative/process models of working memory store verbal information specifically and between the models generally, although some researchers are considering such an attempt (A.D. Baddeley, personal communication, April 19, 1993). So, at a theoretical level, presently there is only a very tentative description of the relationship between the two models of working memory.

There has been not much attempt at an empirical integration of these two models of working memory either. In the present context, there has been little attempt to examine how verbal information is stored in both models. Empirically, the present dissertation operationalises the qualitative/structural model by a verbal simple-span task and the phonological similarity and word-length effects present in that task. The quantitative/process model of working memory is operationally defined by a complex-span task. Thus, one hypothesis that might be verified if the two models of working memory store verbal information similarly is that qualitative effects (phonological similarity and word-length effects) might be equally present in both simple and complex-span tasks. Before considering this hypothesis however, there is an issue with the operational definition of the qualitative/structural model that needs to be considered.

It is possible to argue that the qualitative/structural model cannot be totally defined by a verbal simple span task alone. The model, as described in Chapter 2, also contains visuospatial and central executive components. However, the operational definition used herein does reflect the model's ability to account for verbal information, the only type of information the quantitative/process model accounts for. Until the relationship between the models regarding verbal

information is clarified, the main type of information processed by the quantitative/process model, the present operational definition is sufficient for the purposes of the present research which is to investigate how the Phonological Loop component of the qualitative/structural model relates to the quantitative/process model.

It could also be argued that, by comparing the phonological loop system of the qualitative/structural model and the quantitative/process model which is presumed to be located in the central executive (as argued above), one is simply comparing two aspects of the same model. While this argument ignores the body of research mentioned above which describes these as two separate models, it is nonetheless a germane hypothesis. However, if these are simply two parts of the same model, then one would not expect to find qualitative effects definitive of one subcomponent of the model (e.g., word length, phonological similarity, and articulatory suppression effects) occurring in another subcomponent. Should this occur, one would be hard pressed to explain how the subcomponents can be considered fractionated when fractionation is predicated on the notion that certain effects occur in one, but not another, subcomponent. Thus, if in empirical comparisons of simple and complex-span tasks effects used to define the phonological loop also occurred in the task supposedly using the central executive, this evidence would imply that these tasks are using the same, not separate components. Thus, in comparing these two similar tasks, the current wisdom implies that one would be unlikely to find word-length or phonological similarity effects in the complex-span task because this task *primarily* uses the central executive. To date there appears to be only one study program that has compared complex-span and simple-span tasks directly. La Pointe and Engle (1990) conducted a series of five experiments in which word length was manipulated within both complex-span and simple-span tasks. In the complex-span task of their first experiment, 80 undergraduates were presented with series of sentences followed by a target word;

series ranged from two to five words in length. The target words were either one or three syllables in length, forming the short and long word-length conditions respectively. The participants read the sentences aloud, then recalled the words in any order. The same participants were also presented with a similar series of words to read and recall; the simple-span measure. La Pointe and Engle's second experiment was exactly the same with the exception that the sentence was replaced by a mathematical operation in the complex-span task. In La Pointe and Engle's first and second experiments, short words were better recalled than long words, and words in simple-span tasks were recalled better than words in complex-span tasks. That is, while recall was higher in a simple-span task than in a complex-span task, both tasks produced a word-length effect.

In their third experiment, participants repeated the procedure for the second experiment except that they also concurrently engaged in articulatory suppression. In this third experiment, despite the use of articulatory suppression, a word-length effect was found. However, there was no difference in the word-length effect nor in overall recall as a function of the nature of the task (simple-span, complex-span). In their fourth and fifth experiments, while La Pointe and Engle again compared task-types, and again found word-length effects in both simple-span and complex-span tasks, they do not report any inferential testing of the relationship between task-type and the word-length effects they found.

The point to be taken from La Pointe and Engle's five experiments is that word-length effects occur in both simple-span and complex-span tasks. If, as asserted by those authors discussed above (Gathercole & Baddeley, 1993; Hasher & Zacks, 1988; Just & Carpenter, 1992), the complex-span task assesses the Central Executive, then, from the results of La Pointe and Engle, one must explain how a word length effect occurs in both tasks. Various views exist on this point, including that the Central Executive utilises the phonological loop.

However, these views all rest on certain assumptions about the Central Executive and about the complex span task used to operationalise it. However, rather than explain how word-length effects might occur in the Central Executive and the Articulatory Rehearsal Process, La Pointe and Engle concluded that complex-span tasks do not possess a special relationship with reading comprehension, relative to simple span tasks. Instead, their results imply that *both simple and complex-span tasks rely on the same fundamental processes, such as the Phonological Loop, for assisting in reading comprehension.*

Conclusions

From the data and literature reviewed in the present section, two things are apparent. First, there has been only minimal consideration of the relationship between how the qualitative/structural and quantitative/process models of working memory account for visually presented verbal information. However, the presumption in the literature is that the quantitative/process model is concerned with the Central Executive. In contrast, the empirical evidence, although based on only five experiments from one research program, suggests that the simple-span and complex-span tasks as operationalisations of the two models of working memory demonstrate some commonality. Thus the empirical evidence is inconsistent with the theoretical notion that the quantitative/process model is subsumed in the Central Executive. The empirical data are more consistent with the quantitative/process model of working memory being related to the Phonological Loop of the qualitative/structural model. Given that there is very scant knowledge about the relationship between both models of working memory, and that the available empirical evidence (especially La Pointe & Engle, 1990) is inconsistent with the theoretical relationship (Gathercole & Baddeley, 1993; Just & Carpenter, 1992), it seems appropriate to directly test the relationship. In conclusion, due to the apparent conflict between empirically based and theoretically derived

conclusions about the relationship between the two working memory models, there is an imperative to clarify and describe the relationship between how qualitative/structural and quantitative/process models of working memory account for storage of verbal information.

The Relationship Between Meta-Memory and Working Memory

While the nature of working memory has been defined using two models, I also suggest in this dissertation that working memory represents a process that persons are 'aware' of. That is, working memory is a cognitive operation that persons can monitor and report on, a process that has been termed meta-memory (Flavell & Wellman, 1977). While there is no evidence bearing directly on the relationship between working memory and meta-memory, it was shown in Chapter 4 that there exists an established relationship between meta-memory and other memory performance. Nonetheless, in examining the relationship between working memory and meta-memory, it must be acknowledged that the hypotheses advanced are primarily speculative.

Having an accurate meta-memory for working memory implies that, to the extent that self-report is veridical, persons might be able to report on their working memory. That is, there may well be a direct relationship between a person's working memory and their meta-memory for that process. The present research tests the hypothesis that persons may be able to monitor the operation of working memory in a direct manner by examining their ability to predict the capacity of their working memory.

Furthermore, given that objective working memory performance responds to changes in phonological similarity and word-length, both theoretically defined as indicating operation of the Phonological Loop's components, the present author chose to test whether persons could predict those effects. That is, the present research examined whether, if persons were given only information about the parameter changes in an upcoming task, they would be able to predict the impact of those

changes on recall. Predicting changes in working memory implies that the pattern of the changes in meta-memory and working memory would be the same, and that there would be a predictive relationship between meta-memory and working memory measures.

In summary, the present dissertation aims to explore the extent that persons are able to monitor their working memory. Furthermore, whether meta-memory for working memory is a general process and whether it, like working memory, responds to parameter changes will be examined. Only when information about whether persons have some on-line meta-memory for their working memory has been gathered can research begin to address the potential responsivity of working memory to the information in meta-memory, an issue of bi-directionality (Nelson & Narens, 1990).

Toward an Integration of Theories and Measurement of Verbal Working Memory

The first reason for integration of how the differing models of working memory account for verbal information was to begin to establish a shared language with common meanings in describing verbal information processing in working memory. That is, integration will allow one to answer questions about whether experimental effects are *sui generis* to working memory or not.

Given that integration of verbal working memory definitions is a valid goal, the second reason for integration is that testing predictions from one model in a task used to index another model can lead to new aspects of the integrated model being discovered or described. That is, integration allows greater development of working memory models.

This dissertation organizes and integrates some aspects of how working memory has been both operationalised and theorized about in terms of verbal information processing. The literature suggests two organizing dimensions used in the operationalisation of working memory: the dichotomy between simple-span and complex-span tasks, and whether assessment is

subjective (e.g., self-report) or objective (e.g., performance measured in the laboratory).

There are also two distinct ways of describing theoretical dimensions of working memory: as a quantitative process consisting of what is remembered (content), the order it is recalled (order), and the speed at which the process operates (processing time); and, as a qualitatively defined set of structures such as the Phonological Store and Articulatory Rehearsal Process which are inferred from the phonological similarity, word-length, and articulatory suppression effects on recall of content. Thus, any integration of the theory and assessment of working memory must involve a linkage at both theoretical and operational levels between previously used aspects of working memory. Specifically, the relationship between dimensions of task-type (simple-span and complex-span), method of report (subjective and objective) quantitative/process aspects (content, order, and processing time), and qualitative/structural aspects (phonological similarity, word-length, and articulatory suppression effects) must be investigated.

SUMMARY

At this point a brief summary of the problem(s) that this dissertation addresses is presented. First, there have been two predominant ways of defining and operationalising how working memory accounts for verbal information. From the literature, it was unclear how these two definitions of working memory related to each other at both theoretical and operational levels. Thus, the first goal of the present research was to unravel the relationships between how the qualitative/structural and quantitative/process models of working memory store visually presented verbal information. Second, the present research investigates the relationship between working memory and meta-memory. In doing so, the present research explicitly tests the idea that persons can accurately monitor and report on their working memory, a question that is apparently untested to date. In summary, a

more complete description, at both theoretical and operational levels, of working memory necessarily involves comparison of both models and between subjective and objective measuring methods.

Experimentation

CHAPTER 6: GENERAL HYPOTHESES AND GENERAL METHOD

Prior to presenting the series of experiments in this dissertation, a short review will be presented of the general hypotheses of the present set of experiments. This review is designed to orient the reader as to the theme of the present dissertation, namely the integration of research into working memory. Second, the General Method will then be presented so that the overall plan in answering the general hypotheses will be apparent.

GENERAL HYPOTHESES: THREE DEFINING POSTULATES

The main hypotheses of this dissertation are that: *the two models of working memory are in fact examining the same construct in terms of verbal information storage and that persons can monitor their working memory.* In stating that the two models refer to the same construct, it is important to note that 'construct' is not a synonym for physical entity. That is, there is no assumption in this dissertation that working memory is, or is not, localized in any particular area of the brain. To be sidetracked into conceptualizing working memory as simply an organic entity or as a structure is to minimize the importance of addressing the more abstract notion of how working memory operates as opposed to what it is.

The next step in developing hypotheses is to present the implications of stating that the two models of working memory store verbal information similarly. This will be done in two stages: at a theoretical level where constructs and factors of those constructs are discussed; and at an operational

level where implementations of factors (variables) are discussed.

Hypotheses at a Theoretical Level

At an abstract level, if construct A is equivalent to construct B, then the factors that make up the set of construct A must also be in the set of construct B. Conversely, the factors that make up the set of construct B must also be in the set of construct A. It is this theorem of set membership that is at the heart of the present dissertation. From this theorem, the first step in developing hypotheses is to identify the factors representing the items in each set.

Defining the factors in the hypotheses

The verbal storage component of the qualitative/structural model has been traditionally described as a set of cognitive structures (Figure 1), an issue that Baddeley (personal communication, April 19, 1993) acknowledges. These structures have been identified as the Phonological Store and the Articulatory Rehearsal Process. These structures are inferred from the phonological similarity and word-length effects plus the impact that articulatory suppression has on those effects (Baddeley, 1986, 1992a, 1992b). Thus, it is sensible to refer to the qualitative/structural model's primary factors as being the phonological similarity, word-length, and articulatory suppression effects. From measurements of these three factors, inferences about the Phonological Loop as a whole can be stated in toto and without reference to any other factors.

The factors of the quantitative/process model of working memory have been described as the maintenance of content by working memory (content), the activation of order information (order), and the time it takes for items to be activated and maintained (processing time).

Finally, there have been two parallel research paradigms examining either working memory or meta-working memory. To

avoid confusion, these two factors have been relabelled: working memory is termed an objective factor; and meta-working memory is termed a subjective factor. Thus, when the objective factor of working memory is used, what is being described is working memory as it is perceived by an external observer; subjective working memory is working memory as it is perceived by the person in whom that process resides.

Producing theoretical postulates to integrate working memory factors

In reviewing the previous empirical work examining working memory, it became apparent that there has been little integration of the factors used to define working memory with each other. This lack of integration has been at both theoretical and operational levels.

Three guiding postulates will be used in this dissertation to generate theoretical links between factors. The way that these postulates were generated was from the abstract mathematical form that if A is equivalent to B, then the members of set A will be in set B and vice versa (see above). The first postulate in the present dissertation is:

1. If both quantitative/process and qualitative/structural conceptualizations of working memory are describing the same construct, then the qualitative defining factors of the qualitative/structural model (the phonological similarity and word-length effects) ought to also be evident in a quantitatively based model;

That is, verbal information storage is hypothesised as occurring similarly in both models. In operational terms, Postulate 1 implies that the phonological similarity and word-length effects used to infer the Phonological Loop structure (Baddeley, 1986, 1992a, 1992b) would be evident in typical complex-span tasks involving an operation and recall component as well as in simple-span tasks (La Pointe & Engle, 1990). Previously, these effects have only been described in the context of simple-span tasks. It was also investigated whether, even if these effects are present in both simple and

complex-span tasks, the *magnitude* of these effects alters with the addition of a concurrent operation.

Examining the empirical basis of Postulate 1 largely determined the overall structure of the present set of studies. An outline of this structure and the logic behind it will now be presented to assist the reader in understanding the context of the present studies (Figure 3). The tenor of Postulate 1 suggested that both phonological similarity and word-length effects may be present in both simple-span and complex-span tasks. As the phonological similarity and word-length effects had been demonstrated many times previously in simple-span tasks, it was decided to first test whether this effect could be replicated in a complex-span task for the phonological similarity (Experiment 1) and word-length (Experiment 2) effects. If only one or neither of the phonological similarity or word-length effects were replicated, logic dictates that the same task be re-examined but without the concurrent operation for the phonological similarity (Experiment 3) and word-length (Experiment 4) effects. If these effects were not replicated in a simple-span task, then an investigation of the conditions necessary to replicate the phonological similarity (Experiment 5) and word-length (Experiment 6) effects would be prudent. Once the conditions for replication of the phonological similarity and word-length effects had been established, both simple-span and complex span tasks could be compared to test whether phonological similarity and word-length effects were equivalent in each, and thus directly test Postulate 1 (Experiment 7). It must also be noted the direct comparison of simple-span and complex-span tasks implicitly requires the presence of phonological similarity and word-length effects in both simple-span and complex-span tasks and so repeats the logical process behind the previous six experiments for Experiment 7.

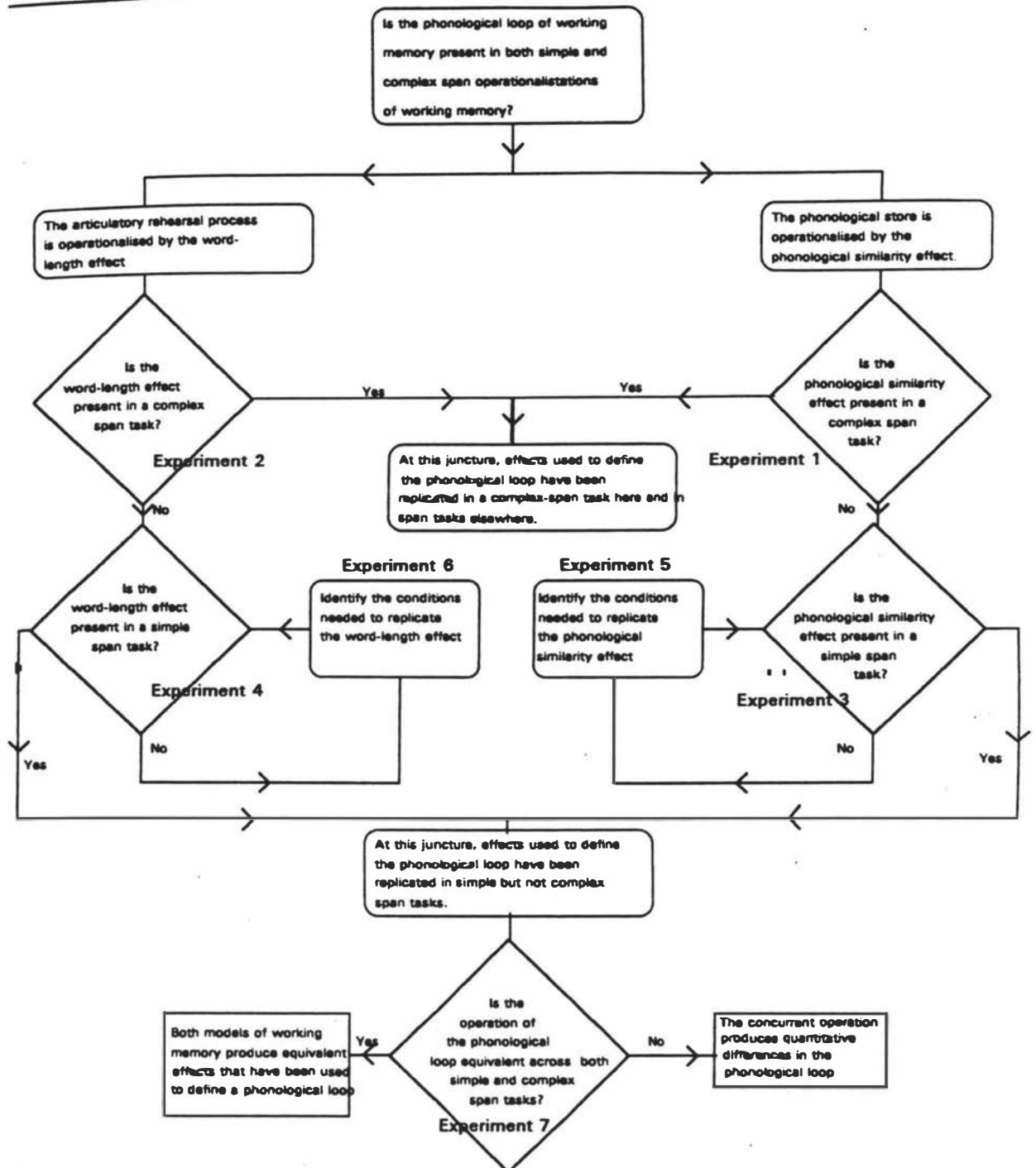


Figure 3. A flowchart showing some aspects of the logic behind the development of the present studies in relation to Postulate 1.

In the present studies, a second postulate was advanced that:

2. The factors of order and processing time of process/quantitative models will be useful in elaborating upon the qualitative/structural model of working memory, in addition to a measure of content.

Operationally, Postulate 2 implies that by also using measures of order and processing time that the phonological similarity, word-length, and articulatory suppression effects (and by implication, the Phonological Loop) could be more exactly described. Typically, content has been the only aspect directly investigated in most previous research on working memory. However, close examination of the measures typically used in studies of working memory indicates that order is frequently confounded with measures of content because items must be recalled serially (e.g., Baddeley, et al., 1975; Baddeley, et al., 1984).

Finally, if working memory can be accurately described by an external observer, then it is worthwhile examining whether the same process may also be accurately described by an internal observer. This produces a third postulate, namely that:

3. The components and not simply the general process of working memory are monitorable by both external and internal observation, thus linking subjective (self-report) and objective (performance) measures of working memory.

Operationally, Postulate 3 implies that content measures of on-line meta-memory (Kausler, 1991) will mirror content measures of memory performance in pattern and will share a relationship with each other. That is, persons were expected to be able to form accurate pre-estimates of their working memory, pre-estimates that reflect effects of phonological similarity, word-length, and articulatory suppression. From Postulate 3, the questions that were posed are as follows: (1) Can persons pre-estimate qualitative changes, such as the phonological similarity, word-length, and articulatory suppression effects, in working memory operation?; and (2) Can persons form accurate pre-estimates of the magnitude of their recall? These two questions suggested that one might predict that the pattern of on-line pre-estimates of meta-memory would be comparable to the performance measures of

content and that there would be a relationship between pre-estimates and performance.

Together, these three sets of general hypotheses from Postulates 1 to 3 oriented the present research toward a method of examining the unity, or not, of working memory and to laying some empirical foundations for investigating the relationship between working memory and meta-memory. These general hypotheses will be instantiated in different ways within each experiment in the present research.

GENERAL METHOD

Participants

Prior to recruiting participants, approval for conducting the present research was obtained from the Massey University Human Ethics Committee (HEC 93/24). All participants were students ranging in age from 16-48 years old. One benefit of using a relatively homogeneous student sample was that some comparison of data across experiments could be performed without explicitly using measures of education level as a covariate. Also, using student samples helped ensure that all participants were relatively familiar with using a computer, thus avoiding a confound of computer literacy (Jay & Willis, 1992; Salisbury, 1990). The experimenter further ensured computer literacy in the sample by asking each person, prior to the experiment, about their computing experience, with the result that all participants reported that they could use a computer adequately. One disadvantage of using a student sample was that the absolute levels of performance may have been artificially high with such a well educated sample. However, as the purpose of the present series of experiments was to compare effects within-groups and not to establish normative data, this disadvantage was not seen as precluding use of a student sample.

No remuneration or other incentives were offered to participants in Experiments 1 to 4. For Experiments 5 and 6, participants were paid \$5 for their participation. For

Experiment 7, participants were either given the option of either \$5 or 2% course credit for their participation. The honorarium was made possible through Massey University Graduate Research Fund Grant A93/G/058 and the course credit through Dr. Julie Bunnell.

Apparatus

The experiments were conducted in a well lit and quiet room with only the participant present for the experimental trials proper.

All of the experimental trials were presented on an Excel IBM-clone personal computer with a low resolution cga screen. Timing of reaction times was facilitated by use of a Turbo Pascal (© Borland) 5.0 *.tpu unit called 'Optimer' which calibrates timing procedure calls to the clock speed of the PC's cpu. The Optimer unit has an accuracy of 1×10^{-3} to 1×10^{-6} s. Typically PC accuracy is in the order of 2×10^{-2} s.

The software used to present and record the experiments was written by the author using Turbo Pascal 5.0 programming language. User-written programming was advantageous in that it allowed careful analysis of the implementation of the experimental procedure into computer code.

Stimuli

Target stimuli were in all instances English (UK) words. Stimuli selection and use will be now described in terms of the classes of stimuli used, pool sampling methods, sentence construction methods, and initial stimuli selection.

Stimuli class

There were four classes of stimuli used over the seven experiments in this dissertation: phonologically similar words (Experiments 1, 3, 5, & 7); 2-syllable words (Experiments 2 and 4); 3-syllable words (Experiments 6 & 7); and a control condition of 1-syllable and phonologically dissimilar words (Experiments 1 to 7; Table 5).

Table 5. A summary of the experimental conditions used in the present dissertation. Conditions vary across phonological similarity, word-length, articulatory suppression, task-type, pool-type, and pacing.

	Experiment						
	1	2	3	4	5	6	7
Phonological similarity	✓	-	✓	-	✓	-	✓
Word-length *	-	✓	-	✓	-	✓	✓
Articulatory suppression	✓	✓	✓	✓	✓	✓	-
Task-type	C	C	S	S	S	S	✓
Stimulus pooltype	-	-	-	-	✓	-	-
Presentation pace	-	-	-	-	-	✓	-

C = stimuli were part of a complex-span task

S = stimuli were part of a simple-span task

*The word length condition compared 1 and 2 syllable words in Experiments 2 and 4, and compared 1 and 3 syllable words in Experiments 6 and 7.

Pool sampling methods

Stimuli (word) lists were of two types in the present series of experiments. The first type, termed throughout as *non-replacement* stimuli, were words drawn randomly and without replacement from a larger word pool. The second type, termed *fixed-pool* stimuli were selected with replacement from a small pool of ten words. Fixed-pool stimuli were randomly formed into trial groups of six items with the condition that no item was repeated in a single trial. Where fixed-pool stimuli were used, the frequency that each target word occurred was approximately equal. Non-replacement stimuli were used in Experiments 1, 2, 3, 4, and 6. Non-replacement and fixed-pool stimuli were used in Experiment 5. Fixed-pool stimuli only were used in Experiment 7 (Table 5). In Experiment 7 in which only sampling with replacement from a fixed pool was used, target words were randomly selected for each trial with two constraints: no words were repeated within any one trial; and each participant received the same set of randomly ordered items (a constant random seed) but with those sets counterbalanced across the four blocks of trials.

Sentence construction

Sentences with target words at the end were used in Experiments 1, 2, and 7 (Table 5). The addition of the

sentence verification task changed the simple-span to a complex-span task.

The sentences used in the complex-span condition in Experiments 1 and 2 were constructed using a subject-predicate form. Target words were always the final word in the sentence (Appendix 1, Table A.1). For example, the stimulus presented to the participant might be *A cat can be blue* where *blue* is the target word to be recalled later. The sentences in the complex-span condition in Experiment 7 were constructed using six lexically-based statement forms with a target word at the end (e.g., *The letter x is in the word **). The lexical task consisted of six sentence forms repeated over each trial in random order (Appendix 1, Table A.6).

For Experiments 1, 2, and 7, the sentences were correct for half the items and incorrect for half the items in each trial. In all instances, sentences were presented in a random order, but with each person receiving the same random order.

Stimuli selection

Stimuli (words) were initially selected from Elley's (1975) word frequency norms. These norms are New Zealand based, and equate nouns for their frequency of usage on a nine point scale. In the phonological similarity and control conditions for Experiments 1 and 2, all words were monosyllabic. The word list of 36 phonologically similar words used in Experiments 1 and 2 had a mean frequency of 5.08 ($SD = 2.27$), and the 36 phonologically dissimilar control words had a mean frequency of 5.00 ($SD = 0$) based on Elley's norms. The 72 target words in Experiments 2 and 4, varying in word-length, had a mean frequency of 5.00 ($SD = 0$; Appendix 1, Table A.1).

For the fixed pool in Experiment 5, the phonologically similar sounding 1-syllable words were constructed from a list used by Brooks and Watkins (1990) and the phonologically dissimilar (control) words were taken from stimuli used in Experiments 1 and 3 (Appendix 1, Table A.2). For the non-replacement pool in Experiment 5, three sets of phonologically similar words were constructed from Elley's

word norms with three further sets being constructed from lists used by Brooks and Watkins (Appendix 1, Table A.3).

The 72 1-syllable and 72 3-syllable words for Experiment 6 were constructed from Elley's (1975) word norms. The 1-syllable and 3-syllable words all had a mean frequency of 5.00 ($SD = 0$; Appendix 1, Table A.4).

Finally, the control and 3-syllable words in Experiment 7 were constructed from Elley's (1975) word norms with a mean frequency of 5.00 ($SD = 0$). The phonologically similar words were constructed from a list by Brookes and Watkins (1990); this was the same list of phonologically similar words as was used in Experiment 5 (Appendix 1, Table A.5).

Equating word frequency across all experiments in the present series helped ensure that some stimuli were not better recalled than others simply because they were more familiar to the person. In addition, Elley's (1975) norms were used in preference to the older Kucera and Francis (1967) norms for two reasons. First, Elley's norms were developed in 1975 from words specifically used in New Zealand. Second, the age of the persons in the study indicated that they were the same cohort for whom the norms were developed. The ages of the persons the norms were developed for ranged from 5 to 15 and up in 1975. The norms were developed from frequency counts of New Zealand school reading journals used in 1975. Thus those person would now be in the age range of 25 to 35. Given that the age of participants in the present experiments was between 16 and 48, the participants are obviously within the cohort for whom the norms were initially developed.

Design

The basic design of Experiments 1 to 7 was to present participants with a task that varied across conditions of word-type, task-type, and articulatory suppression; these were all within-persons conditions where they were used. Only in Experiments 5 and 6 were there any between-subjects conditions. These were pool sampling method (Experiment 5) and stimulus presentation pace (Experiment 6).

Word-type

In each of Experiments 1 to 7, trials were presented to participants in two word conditions: a control condition of 1-syllable dissimilar words (e.g., *sock, trick, sale,...*) and an experimental condition. For Experiments 1, 3, and 5, the experimental condition used phonologically similar words (e.g., *bug, rug, hug,...*). For Experiments 2 and 4 the experimental condition involved 2-syllable words (e.g., *contest, garage, hobby,...*). For Experiment 6, the experimental condition involved 3-syllable words (e.g., *basketball, settlement, chocolate,...*). Finally, for Experiment 7, there were two experimental conditions: phonologically similar words and 3-syllable words (Table 5).

Task-type

There were two possible task-types in Experiments 1 to 7. The first task-type was the simple-span task which involved presentation of a word alone. The second task-type was the complex-span task which involved presentation of a word as part of a sentence that had to be verified as correct or muddled. A complex-span task was used in Experiments 1 and 2. A simple-span task was used in Experiments 3, 4, 5, and 6. Experiment 7 directly compared simple-span with complex-span tasks and so used both task-types within subjects (Table 5).

Articulatory suppression

In Experiments 1 to 6, participants were required to perform half of the experimental trials while performing an articulatory suppression task. Articulatory suppression consisted of repeatedly articulating the digits 1 to 8 in numerical order, at the rate of one digit per second. The articulatory suppression task was implemented to prevent sub-vocalization of items (Murray, 1967, 1968, Murray et al., 1988) and to block the operation of the Articulatory Rehearsal Process. Baddeley, et al. (1984) had also found that articulatory suppression was most effective when used not only during the reading of the stimuli, but during recall

also. Consequently, the present studies used articulatory suppression during both stimuli presentation and recall phases (Table 5).

Procedure

Each experiment consisted of a series of general steps. Upon arrival, each person was greeted and provided with both an information sheet (Appendix 3) and an informed consent sheet (Appendix 4). Each participant was able to ask any general questions about the design at this point.

Next, each participant was seated at the computer that was to present the stimuli. The computer prompted each participant for some basic demographic information (name, participant identification number, age, and gender). This time also served to allow participants to familiarize themselves with the operation of the computer.

Following the gathering of demographic information, each person was presented with an interactive series of computer screens with step-by-step instructions on the upcoming task.

Following the description of the task, each participant was then given four practice trials of the actual task. These trials allowed each person to familiarize themselves with the task and to begin to form some expectations of their performance. No feedback about their performance on the practice trials was ever given so as to avoid confounding the pre- or post-estimates they made of their performance.

Single trial procedure

Prior to each trial, each person was asked to enter a pre-estimate (from 0 to 6) of how many words they thought they would recall in any serial position. These data were used to calculate the pre-estimate variable representing meta-memory monitoring.

Each trial consisted of presentation of a series of six stimuli consecutively. The trial size of six stimuli was determined from a review of the literature and from some

pilot testing as being optimal in avoiding either ceiling or floor effects for a highly educated sample. Stimuli were either a word alone, or a word at the end of, and part of, a sentence. For example, participants would see either the word *horse* or a sentence such as *A cat can ride a horse*. Where sentences were presented for verification, in all instances there were three that were to be verified as correct and three as muddled. The order of the sentences was randomly allocated to the trial.

When the word was presented alone, the person had only to read (silently) the word and then press the space-bar to view the next word. When a word was presented at the end of a sentence, the person had to read the sentence (silently) then verify whether the sentence was correct (by pressing the 'C' key) or muddled (by pressing the 'M' key). A 100 ms interval was provided between the pressing of the response key and onset of the next item. A word presented alone was defined as a *simple-span task*, a word at the end of a sentence was defined as a *complex-span task*. In all instances, the time that the person viewed each stimulus was recorded by the computer to 1 millisecond of accuracy. The only departure from this general procedure was for Experiment 6 in which one group of persons were presented with stimulus items at a fixed pace of one item per second. All reaction time data were used to calculate the viewing and verification time measures of processing time.

After the participant had viewed the six stimuli, the computer emitted a tone to indicate that there were no more stimuli to be presented. Immediately after the tone, a screen was presented asking each person to write all of the words that they recalled onto a sheet of paper. The person was requested to write as many of the words as they could in the correct serial position. The response sheet was constructed so that there were six horizontally arranged squares labelled from 1 to 6 to indicate serial position. If the person could not recall the serial position of the item but could recall

what the item was they were instructed to guess the position of the item.

For Experiments 1 to 4, there was no time limit on how long the person had to recall stimuli. For Experiments 5 and 6, the person had 20 seconds and for Experiment 7, the person had 15 seconds to recall all the items they could remember. To equate writing time within subjects, each person was instructed to write only the first four letters of each item.

Presentation order of word-type, task-type, and articulatory suppression

Each person who completed Experiments 1 to 6 was presented with the same order of experimental blocks. There were a total of 12 trials per session in Experiments 1 and 3. The first six trials were conducted without articulatory suppression (the control condition), and the last six trials with articulatory suppression. Within each set of six trials, phonologically similar words were used for the first three trials and phonologically dissimilar words were used for the last three trials. There were also 12 trials per session in Experiments 2 and 4. Again, the first six trials were conducted without articulatory suppression (the control condition), and the last six trials with articulatory suppression. Within each set of six trials, monosyllabic words were used for the first three trials and two-syllable words were used for the last three trials.

This procedure was repeated over two sessions for Experiments 1 and 2 and one session each for Experiments 3 and 4. The two sessions of Experiments 1 and 2 were separated by a two week interval. Thus Experiments 1 and 2 presented 24 trials across the four conditions whereas Experiments 3 and 4 presented a total 12 trials across the four conditions. Experiments 5 and 6 had the same order of trial presentation as Experiments 1 and 2, respectively, except that trials 13 to 24 were presented immediately after trials 1 to 12 in Experiments 5 and 6.

Summary of the general procedure

The general procedure of Experiments 1 to 7 can be summarized as follows:

1. Arrival and completion of informed consent.
2. Gathering of demographic information.
3. Explanation of the experimental procedure.
4. Presentation of practice trials.
5. Presentation of experimental trials:
 - 5.1 Measurement of the person's pre-estimate.
 - 5.2 Presentation of six words, one at a time.
 - 5.3 Prompt to recall the target words.
 - 5.4 Post-estimate of recall (Experiments 5 to 7 only).
 - 5.5 Inter-trial break of 100 ms.
6. Measurement of the person's reading rate for each word-type.

Dependent variables

The dependent variables for all experiments in the present series were the same. There were five classes of dependent variables: a meta-memory score, two content scores, an order score, a processing time score, and an articulation rate score.

Meta-memory

Meta-memory was calculated as the mean number of items a person predicted that they would recall. This was termed an on-line pre-estimate of their performance. Pre-estimates were taken prior to the trial.

Content

From the number of items recalled in the correct serial position in which they were presented a mean score of recall *in the correct serial position* was calculated. From the total number of items recalled, without regard for which serial position they were presented in, a mean recall *in any serial position* score was calculated for each condition.

Following the recall phase for Experiments 5 to 7, a prompt appeared asking the person to estimate how many items they had recalled in the correct position. This post-estimate of recall was implemented to examine if persons were guessing many items in the correct serial position. If post-estimates were significantly lower than recall in the correct serial

position, this would indicate that serial position was being guessed correctly often.

Together, recall in any serial position, recall in the correct serial position, and post-estimates of recall in the correct serial position were used to operationalise the content factor of working memory.

Order

Experiments 1, 2 and 7 used a variant of the Reading Span task as described by Daneman and Carpenter (1980). The complex-span task used in the present studies differed from that used by Daneman and Carpenter in that the present version presented a fixed-length list of six items to each person, a length chosen to be higher than a typical span. The reason a fixed list-length was used in the present studies was to enable calculation of the number of order errors. Calculation of order errors can only be done when recall in the correct serial position and recall in any serial position scores are available. Order errors were calculated as the difference between recall in any serial position and recall in the correct serial position. The mean number of order errors was used to represent the order factor of working memory for each person.

Processing time

For each trial, the time between presentation of the target item and when the person either (a) verified the sentence in the complex-span task, or (b) pressed the space-bar for the next item in the simple-span task, was recorded to 1 millisecond accuracy. These times were averaged for each condition for each participant and were termed the viewing time score.

Articulation rate: a measure of the Articulatory Rehearsal Process

For Experiments 5 to 7, a measure was taken of how rapidly the person could read the target words aloud. At the end of the experimental trials, a list of all the words used in each word-type condition was presented. The person initiated the list by pressing the space-bar. The person then had to read

the words aloud as rapidly as possible. At the end of the list, the space-bar was re-pressed, stopping the timer. For each list a reading time was measured to 1 millisecond accuracy. The reading times were used to calculate mean articulation rate for each word-type condition.

Analysis

All of the data in the present dissertation were analyzed using SPSS PC+, Version 3.0. In all *a priori* statistical tests the MANOVA procedure was used and the independent variables of task-type, word-type, and articulatory suppression were declared as within-subjects factors using the /WSFACTORS command. Where mixed designs of between-subjects and within-subjects independent variables were used, inferential data are reported as averaged tests of significance.

In all instances for Experiments 1 to 6, *a priori* testing consisted of an examination of univariate main effects followed by examination of interaction terms. Where between subjects factors were used (Experiments 5 and 6; pooltype and pacing, respectively), the main and interaction effects reported are averaged tests of significance. For Experiment 7, *a priori* testing involved separate tests of the one control group against the phonologically similar or word length conditions. Otherwise, testing did not differ from the regimen for the previous six experiments. In all of the experiments the F-distribution was used to provide estimates of the probability levels associated with the F and error terms.

Post hoc testing in all of the present experiments consisted of paired t-tests without corrections for significance when comparing individual cells.

Finally, regression analyses were simple regressions with only a single variable entered against the dependent variable. Where variables have been transformed, this is indicated.

The decision level for all inferential statistics was at $p < .05$ in keeping with psychological conventions (Keppel, 1982). In the dissertation itself, exact significance levels are reported irrespective of whether *a priori* or post-hoc inferential testing was performed. The benefit of reporting exact significance levels is that the reader has the opportunity of applying whatever corrections to the decision level that her or she chooses.

CHAPTER 7: EXPERIMENTS 1 & 2: AN INITIAL EXAMINATION OF QUANTITATIVE/PROCESS, QUALITATIVE/STRUCTURAL, AND SUBJECTIVE/OBJECTIVE FACTORS IN A COMPLEX-SPAN TASK

The first two experiments in this dissertation were an initial test of Postulates 1 to 3. The aim of Experiments 1 and 2 was to examine whether the defining effects of the qualitative/structural model could be replicated in a complex-span task; to explore the utility of assessing dimensions of content, order, and processing time; and to examine the ability of persons to predict their own working memory operation.

The complex-span task used in the present dissertation was a variant of the Reading Span task (Daneman & Carpenter, 1980, 1983). Baddeley (1990) has suggested that the Reading Span task probably involves "...strategy selection, the phonological loop, and a knowledge of vocabulary as well as the capacity to co-ordinate these various aspects of memory" (p.138). In short, the Reading Span task may involve both the Phonological Loop and the Central Executive. In previous studies of the Phonological Loop, both word-length and phonological similarity effects have been found in subsequent recall of words or letters (Baddeley, et al., 1984; Baddeley, et al., 1975; Conrad & Hull, 1964; Vallar & Baddeley, 1984). If, as suggested by Baddeley, the Reading Span task involves the Phonological Loop, then it would be reasonable to expect both the phonological similarity and word-length effects to be present in subsequent recall of words presented in the present variant of the Reading Span task.

Second, the first two experiments examined whether it is useful to utilize measures of content, order and processing time when studying working memory. Salthouse has shown, in his consideration of aging issues, that it is conceptually helpful to examine working memory as a set of dimensions or

factors (e.g., Salthouse, 1985, 1988a, 1988b, 1988c, 1990, 1991a, 1991b, 1992, 1993a, 1993b). Conceptualizing working memory as a set of dimensions allows an extension of more traditional research designs in which measures of what was recalled have been used, to also include measures of order of recall and of the temporal limitations of working memory. That is, in the present study, working memory was conceptualized as having content, order, and processing time dimensions. Multi-dimensional constructions of working memory allow questions to be asked about the operation of working memory as well as questions about the capacity and/or content of working memory.

Previously, only in the areas of applied cognition has a multi-dimensional working memory been examined in terms of content, order, or processing time. For example, Daneman and Carpenter (1980) examined working memory in relation to reading, Salthouse and colleagues have examined working memory in terms of differential changes with age (e.g., Salthouse, 1993a, 1993b; Salthouse & Babcock, 1991; Salthouse, et al., 1991). What this dissertation does is examine working memory and its dimensions with reference to theory about working memory instead of theory about the relationship of working memory to some other applied area. The benefit of focussing solely on working memory is that the present research is not distracted into reifying working memory into a construct that explains a second construct without first defining exactly what working memory is. In summary, Experiments 1 and 2 both examined areas of content, order, and processing time to help provide some empirical evidence in support of a multi-dimensional construction of working memory, as described by Postulate 2 (p. 88).

Persons' ability to predict the operation of their own working memory was also examined in these first two experiments. Examining the ability of persons to predict the

operation of their own working memory was largely investigatory in these first two experiments.

In addition, Experiments 1 and 2 in the present dissertation examined the reliability of performance in the complex-span task by testing participants on two occasions separated by an interval of two weeks (Anastasi, 1982). If working memory scores represent a relatively stable cognitive process, then both scores and changes in scores over time can be inferred to indicate both the state of and changes in working memory.

Having outlined the broad goals of Experiments 1 and 2, largely investigatory goals, it is now appropriate to operationalise those goals in terms of the phonological similarity (Experiment 1) and the word-length effects (Experiment 2).

EXPERIMENT 1: AN INITIAL TEST OF THE PHONOLOGICAL SIMILARITY EFFECT IN A COMPLEX-SPAN TASK

Experiment 1 examined whether the phonological similarity effect could be replicated in a complex-span task; whether content, order, and processing time measures of phonological similarity contributed differential information; and whether performance was stable over time.

Postulate 1: The phonological similarity effect in a complex-span task

Experiment 1 investigated the hypothesis derived from Postulate 1 (p. 85) that the phonological similarity effect would be present in a complex-span task. The phonological similarity effect is when fewer words that sound similar are recalled in the correct serial position than words that are either semantically similar or that are phonologically dissimilar (Baddeley, 1966a, 1966b; Conrad & Hull, 1964; Kintsch & Bushke, 1969). The phonological similarity effect occurs for both auditorily and visually presented words and letters (e.g., Baddeley, et al., 1984). However, when

articulatory suppression is invoked, auditorily presented words show the phonological similarity effect but visually presented words do not. Visually presented words, it has been suggested, must be transformed by the Articulatory Rehearsal Process before entering the Phonological Store (Baddeley, 1986, 1992a, 1992b). When the Articulatory Rehearsal Process is blocked by articulatory suppression, it has been concluded that visually presented words are prevented from being encoded into the Phonological Store (Richardson, et al., 1980; Salame & Baddeley, 1982). Hence, the phonological similarity effect disappears. The selective effect of articulatory suppression on visually presented words has been proposed as due to auditorily presented words having obligatory access to the Phonological Store, whereas visually presented words do not.

The present variant of the Reading Span task used visual presentation for all words. If the present complex-span task involves the phonological loop, then phonological similarity effects should be observed when articulatory suppression was not used. In order to replicate previous demonstrations of the phonological similarity effect (Baddeley, 1966a, 1966b; Conrad & Hull, 1964), three hypotheses were posited for Experiment 1:

- 1.1 There would be greater recall of phonologically dissimilar than similar items;
- 1.2 There would be lower recall of all items when the person is using articulatory suppression; and that
- 1.3 There would be no phonological similarity effect with concurrent articulatory suppression.

Postulate 2: Differentiation of the phonological similarity effect using measures of order and processing time

In addition to investigating how phonological similarity affected recall of words in their correct serial position (content measures), the present experiment also extended previous investigations by specifically addressing the changes in order errors and stimulus viewing times as a function of word-type and articulatory suppression, as suggested by Postulate 2 (p. 88).

Order errors

One goal of Experiment 1 was to determine whether recall in any serial position provided parallel information to the estimate of recall in the correct serial position, or whether the effect of order contributed some additional information about working memory in general and the phonological similarity effect in particular. If phonological similarity effects were found to be linked to order errors, then it could be suggested that the Phonological Store is where order information is maintained. Alternately, if order errors were found to vary as a function of articulatory suppression, then it could be advanced that the Articulatory Rehearsal Process also codes the order of those words. If order errors were independent of articulatory suppression effects, then it could be concluded that the Articulatory Rehearsal Process maintains word content but not order information. Order errors were measured by comparing recall of words in any serial position with recall of words in the correct serial position.

Processing time

Processing time assesses the mean time spent viewing and verifying each stimulus. While it is impossible to know exactly what else occurs during viewing time, it would seem reasonable to assume that at least three things occur. First, the person must read and comprehend the sentence. Second, the person must decide whether the sentence is correct or muddled

and respond accordingly. And third, the person must attempt to preserve a trace of the target (final) word in the sentence. Thus, any change in mean viewing time might represent a change in one or more of these processes. It seems unlikely that viewing times will change because of a change in reading speed, or that verification speed will be affected by the word-type condition. This leaves the rehearsal time for each word as the only factor to change the viewing times. So, in Experiment 1, if similar words are less well remembered than dissimilar words, this may be solely due to less time spent rehearsing those words. This would be confirmed by shorter verification times for phonologically similar than dissimilar items. If, however, recall is different for similar and dissimilar words but viewing times are equivalent, then the data used by Baddeley to infer a Phonological Store will have been replicated here (Baddeley, 1986, 1992a, 1992b).

In contrast, articulatory suppression may interrupt all three processes operating at input by forcing some form of time-sharing of the reading, sentence verification, and rehearsal aspects with digit articulation. Consequently, it might be expected that with articulatory suppression there would be longer viewing times than without articulatory suppression. If viewing times are not increased by articulatory suppression, one could suggest that the Articulatory Rehearsal Process operates independently of the input processes.

Operationally, Postulate 2 implied that by using measures of order (order errors), and processing time (verification and viewing time) for each target, that the phonological similarity and articulatory suppression effects, and by implication, the Phonological Loop, could be more exactly described than by using a combined measure of content and order alone. There were no specific hypotheses that could be derived for Postulate 2 from the literature. Thus Experiment

1 represented an original investigation of the separate aspects of content, order and processing time in relation to the phonological similarity effect.

Postulate 3: Prediction of the phonological similarity effect through monitoring of working memory

From Postulate 3 (p. 88) it was also suggested that persons might be able to predict their own performance as a function of both phonological similarity and of articulatory suppression. If the main and interaction effects of word-type and articulatory suppression were able to be accurately predicted in either relative or in absolute terms, then there would be some basis for speculating that the phonological similarity effect might be able to be monitored.

Operationally, Postulate 3 implied that content measures of on-line meta-memory (Kausler, 1991) would mirror content measures of memory performance in direction and additionally demonstrate some shared variability with performance measures. Again, it could not be hypothesized whether persons would be able to predict the phonological similarity or articulatory suppression effects as no previous work had been published on this question.

Method

Participants

Fourteen female participants aged 19 - 42 (Mean = 25.3 years; SD = 6.2) and 7 male participants aged 19 - 28 (Mean = 22.9; SD = 3.6) completed Experiment 1, a total of 21 participants in all. Participants were recruited from undergraduate psychology classes.

Design

Experiment 1 was a three-way factorial design, with all factors varied within subjects. The first factor was word-type (phonologically similar or phonologically dissimilar), the second factor was articulatory suppression (control or present), and the third factor was sentence type (correct or muddled). Experiment 1 was conducted in two sessions separated by an interval of two weeks. The dependent variables were as detailed in the General Method section (p. 99).

Procedure

The general procedure followed the format described in the General Method section (p. 95).

Results

Prior to presenting the results of Experiment 1, the issues of score reliability across sessions and the sensitivity and bias of persons to the sentence verification task will be presented. Next, measures of content, order, processing time, and on-line meta-memory will be presented. The results generally failed to show a phonological similarity effect. The utility of measures of content, order, and processing time in defining working memory appeared to be high. Finally, on-line meta-memory accuracy was related to performance. Thus, the following results supported Postulates 2 and 3 (p. 88), but only partially supported Postulate 1 (p. 85).

Reliability, sensitivity, and bias

Prior to making any inferences about the complex-span task as a measure of working memory, it is necessary to demonstrate that scores on the complex-span task remain stable over time and that there are no practice effects over that interval.

Stability over time was calculated using the Pearson product-moment correlation coefficient. Table 6 shows the correlation coefficients for each dependent variable as a function of articulatory suppression (control or present) and word-type (phonologically similar or phonologically dissimilar) conditions. For recall in the correct serial position, there were significant correlations across the two week interval for both word-type conditions in the control (articulatory suppression absent) condition but not in the articulatory suppression condition. Recall in any serial position showed significant correlations only for phonologically dissimilar words over the two week interval. Finally, although the verification time and pre-estimate measures were both without two participants' data for the second session (reducing the sample size from $N = 21$ to $N = 19$) both those dependent variables showed moderate to strong correlation coefficients over the two week interval. In conclusion, recall in the correct serial position and recall in any serial position showed some test-retest reliability for some conditions. In contrast, measures of verification time and pre-estimates of recall both showed moderate to high test-retest reliability. From these data, there is some evidence that measures of working memory remain stable over short time periods, even with small sample sizes and only three trials per session.

Furthermore, on the dependent variables of recall in the correct serial position, recall in any serial position, order errors, verification time, and pre-estimates, there were no main effects between sessions when tested with the MANOVA design¹. That is, there were no practice effects across sessions of any theoretical importance. Consequently, the

¹ There were only two interactions between session and any of the other variables; there was a session by suppression by word-type interaction ($F(1,37) = 5.21, p < .05, MSe = 0.69$) and a session \times verification \times word-type \times suppression interaction ($F(1,37) = 8.56, p < .01, MSe = 1.04$) for recall in any serial position. As removing the effect of session did not alter any of the conclusions made from the analysis of recall in any serial position, these two interactions are viewed as simply being statistical aberrations with no relevance to the points being made in this dissertation.

data for both sessions were collapsed across sessions using the AGGREGATE command in SPSS.PC+. Aggregation of data across sessions had the added advantage of making each person's score for each condition out of 6 trials, instead of out of three trials for each session. In summary, data from the present complex-span task showed some stability over time.

Table 6. A summary of the two week test-retest reliability of the complex-span task for phonologically similar and dissimilar words with and without articulatory suppression in Experiment 1. Test-retest reliability was calculated using the Pearson product-moment coefficient.

	Recall in the correct serial position ^{*1}	Recall in any serial position ^{*1}	Verification time ^{*2}	Pre-estimates ^{*2}
Control				
Similar	.55**	.30	.67**	.67 **
Dissimilar	.64**	.48*	.62**	.78 ***
Articulatory suppression				
Similar	.22	.40	.83***	.65 **
Dissimilar	.33	.44*	.86***	.74 ***

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

*1 N = 21

*2 N = 18

It could also be argued that how accurately participants responded to the sentence verification task and whether they were more or less likely to respond correct or muddled might have confounded the other dependent variables. The propensity of persons to respond accurately and to respond equally as frequently that a sentence was correct or muddled are issues of sensitivity and bias respectively. However, a signal detection analysis of participants' sensitivity and bias showed that, in Experiment 1, participants had very high sensitivity and responded without any bias (Appendix 2; Table A.7). That is, participants were able to correctly detect whether a sentence was correct or muddled and to respond without bias.

Measures of Content

There were two operationally defined measures of content in the present study: recall in the correct serial position and recall in any serial position. Although it could be argued that recall in the correct serial position is inherently confounded with order, its use is conventional in studies of working memory. Because one aim of Experiment 1 was replication of previous findings, use of the ordered recall measure was necessary.

Recall in the correct serial position

For recall of words in the correct serial position (Figure 4) there was no phonological similarity effect. In fact, phonologically similar words were better recalled than phonologically dissimilar words overall, $F(1,20) = 5.65$, $p = .028$, $MSE = .22$. Thus Hypothesis 1.1 was not accepted.

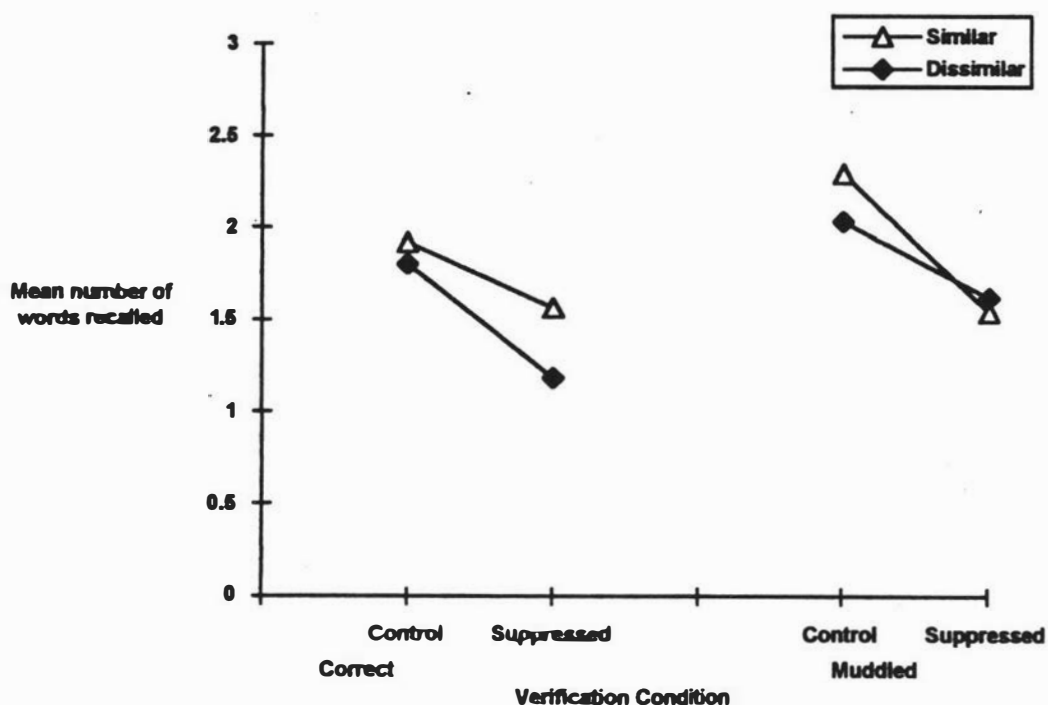


Figure 4. Recall in the correct serial position by phonological similarity, articulatory suppression, and verification conditions for Experiment 1.

As predicted by Hypothesis 1.2, articulatory suppression reduced the level of target word recall overall, $F(1,20) = 64.61$, $p < .001$, $MSE = .19$. Whether words were verified as correct or muddled also affected recall, $F(1,20) = 18.54$, $p < .001$, $MSE = .15$, such that words in sentences verified as muddled were better recalled in all but the phonologically similar and suppressed trials than words in sentences verified as correct (Table 7).

Because there was no phonological similarity effect, no comment as to an interaction predicted by Hypothesis 1.3 between word-type and articulatory suppression was able to be made.

The effect of verifying a sentence on recall in the correct serial position

There was a significant sentence verification \times word-type \times articulatory suppression condition interaction, $F(1,20) = 8.40$, $p = .009$, $MSE = .11$. While the interaction is of no particular importance to the hypotheses in the present research, it is worthwhile to explore the effect of verifying a sentence as correct or muddled, if only as a manipulation check. Consequently, the nature of this three-way interaction will be explored below by considering the correct and muddled verification conditions separately.

In the correct verification condition, articulatory suppression lowered overall recall, $F(1,20) = 22.65$, $p < .001$, $MSE = .22$. There was also higher recall of phonologically similar than dissimilar words, $F(1,20) = 12.35$, $p = .002$, $MSE = .11$. There was no interaction between articulatory suppression and word-type. Paired t-tests were used to compare individual conditions across word-type. There was a difference in word-type with articulatory suppression, $t(20) = 4.34$, $p < .001$, two-tailed, but there was no word-type difference without articulatory suppression (Figure 4).

In the muddled verification condition, articulatory suppression again lowered overall recall, $F(1,20) = 74.46$, p

$< .001$. $MSE = .10$. Unlike the correct condition, there was equivalent recall of phonologically similar and dissimilar words for the muddled verification condition. Again, there was no interaction between articulatory suppression and word-type. Paired t-tests were again used to compare individual conditions across word-type. There were no differences in recall across word-type for either suppression condition (Figure 4).

In conclusion, the three-way interaction was due to the strong difference between word-types that only occurred in the correct verification with articulatory suppression cell. As stated before, it is unclear what, if any, theoretical importance this result has for the present research.

Recall in any serial position

For recall in any serial position, the results were essentially replications of the results of recall in the correct serial position (Figure 5). There was lower recall of phonologically dissimilar than phonologically similar words, $F(1,20) = 120.76$, $p < .001$, $MSE = .12$, contravening Hypothesis 1.1. There was lower recall in the articulatory suppression condition than in the control condition, $F(1,20) = 120.76$, $p < .001$, $MSE = .12$, confirming Hypothesis 1.2.

There was no interaction between word-type and articulatory suppression. Finally, there were no main effects of verification condition. Because there were no main effects of the verification condition and because the interactions between verification and the other conditions add no information to the theme of this dissertation, they will not be reported here.

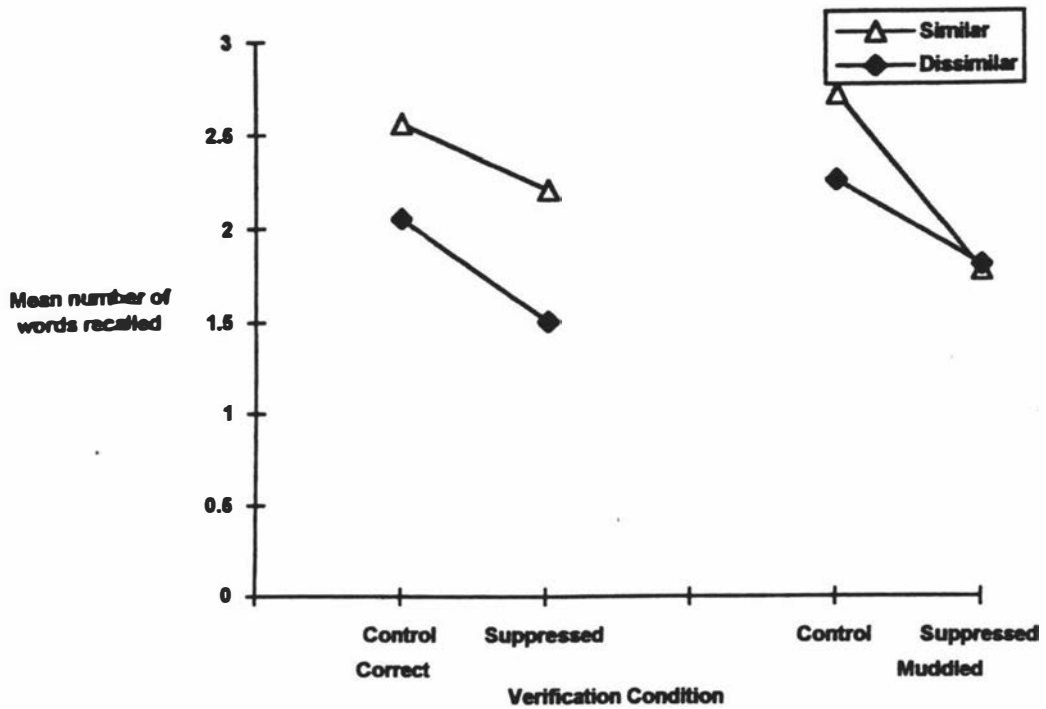


Figure 5. Recall in any serial position by phonological similarity, articulatory suppression, and verification conditions for Experiment 1.

Summary of the results for content

To summarise, on both dependent variables of content, phonologically similar words were better recalled than phonologically dissimilar words. This is the opposite to what was predicted by Hypothesis 1.1. Consequently, there was no phonological similarity effect in Experiment 1. The unexpected better recall of phonologically similar than dissimilar words will be discussed in more depth in the Discussion section of Experiment 1.

As predicted, articulatory suppression lowered the overall recall of words irrespective of which measure of content was used. Thus Hypothesis 1.2, which was based on the premise that articulatory suppression would impair the Articulatory Rehearsal Process responsible for maintaining content in the Phonological Loop, was supported.

Without a phonological similarity effect, there can obviously be no comment made about whether articulatory suppression

removes the phonological similarity effect. Thus, the status of Hypothesis 1.3 must remain as indeterminable from Experiment 1, and it is unclear what effect, if any, articulatory suppression in a complex-span task might have on the phonological similarity effect.

Table 7. Mean (SD) recall of phonologically similar or phonologically dissimilar words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions in Experiment 1. Note that recall is out of three for either correct or muddled verification conditions and out of 6 for the total verification and pre-estimates columns (N = 21).

	Recall in the correct serial position			Recall in any serial position		
	Correct	Muddled	Total	Correct	Muddled	Total
Control						
Similar	1.92 (0.62)	2.29 (0.47)	4.21 (1.00)	2.56 (0.32)	2.71 (0.29)	5.28 (0.48)
Dissimilar	1.80 (0.80)	2.03 (0.67)	3.83 (1.23)	2.06 (0.47)	2.26 (0.53)	4.32 (0.89)
Articulatory suppression						
Similar	1.56 (0.50)	1.54 (0.52)	3.10 (0.86)	2.21 (0.39)	1.78 (0.53)	3.98 (0.70)
Dissimilar	1.18 (0.58)	1.62 (0.60)	2.80 (1.09)	1.50 (0.44)	1.81 (0.48)	3.31 (0.83)

Finally, a manipulation check was performed on whether the operation of the phonological similarity and articulatory suppression effects differed depending on whether the sentences in the complex-span task were correct or muddled. There was no reason from these analyses to expect that the operation of either of these effects would be different depending on whether the sentence was correct or muddled, and no differences were observed.

The effect of order

The effect of order was calculated as the number of words recalled in any serial position minus the number of words recalled in their correct serial position (Table 8). While this is not a perfect measure of the number of order errors, it is a very strict definition which allows no subjective judgment of whether an order error occurred or not.

There was no practice effect evident from the first to the second session, so data were aggregated from both sessions as

was done for the content measures. There were more order errors for phonologically similar than dissimilar words, $F(1,20) = 51.89$, $p < .001$, $MSE = .05$. Articulatory suppression had no effect on the number of order errors, and there was no interaction between word-type and articulatory suppression.

Table 8. Mean (SD) order errors for phonologically similar or phonologically dissimilar words for both control and articulatory suppression conditions in Experiment 1. Order errors are out of 3 for the correct and muddled conditions, and out of 6 overall ($N = 21$).

	Mean order errors (SD)		
	Correct	Muddled	Total
Control			
Similar	0.64 (0.53)	0.42 (0.42)	1.06 (0.76)
Dissimilar	0.25 (0.35)	0.23 (0.26)	0.48 (0.51)
Articulatory suppression			
Similar	0.64 (0.25)	0.24 (0.36)	0.88 (0.43)
Dissimilar	0.32 (0.36)	0.19 (0.30)	0.51 (0.51)

The effect of verifying a sentence on order errors

Sentences verified as correct had more order errors than those verified as muddled, $F(1,20) = 10.11$, $p = .005$, $MSE = .16$. There was also a verification \times word-type interaction, $F(1,20) = 8.78$, $p = .008$, $MSE = .07$. This interaction was due to there being more order errors for phonologically similar than dissimilar items in the correct verification condition, $F(1,20) = 65.22$, $p < .001$, $MSE = .04$, but not in the muddled condition.

Summary of the results for order

To summarize, the effect of requiring words to be recalled in the correct serial position was to lower recall relative to recall in any serial position, indicating that recalling items in order requires extra resources of some type. Because articulatory suppression did not influence the number of order errors, it can also be tentatively suggested that it is not the Articulatory Rehearsal Process that stores word order information. Furthermore, it also appeared that the order of

phonologically similar words was less easily recalled than phonologically dissimilar words, suggesting that word similarity is linked in some way with the storage of order information.

It was also found that sentence verification processes were related to storage of word order information because of the greater effect of correct verifications on order than of muddled verifications. It is unclear if there is any theoretical significance to this finding other than the need to have equal correct and muddled verifications in each trial or block of a research design.

Processing time: The impact of verification and viewing time on recall

The present research found that sentences ending in phonologically similar words were verified more slowly than sentences ending in dissimilar words, $F(1,20) = 17.67$, $p < .001$, $MSE = 140,594$. These are the same words for which recall was better, raising the possibility that the more time spent verifying a sentence, the greater the subsequent recall of the final word. The correct condition was slightly slower (68 ms) in mean verification time when compared with the muddled condition, $F(1,20) = 5.20$, $p = .034$, $MSE = 38,060$ (Table 9). Finally, an interaction between articulatory suppression and verification conditions was found, $F(1,20) = 11.87$, $p = .003$, $MSE = 168,051$, such that in the muddled condition articulatory suppression significantly reduced time spent verifying the sentences, $F(1,20) = 8.39$, $p = .009$, $MSE = 125,587$, but had little effect in the correct verification condition (Table 9).

The main result of Experiment 1 in relation to viewing time of each stimulus was that more time was spent viewing and verifying sentences containing phonologically similar than dissimilar words. However, when viewing and verification time was regressed against recall in the correct serial position, there was no causal relationship between the two variables.

That is, there is no basis from the present data to assert that viewing and verification time predicts recall.

Table 9. Mean (SD) verification times (ms) for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 1 (N = 21).

	Verification times (ms)		
	Correct	Muddled	Mean
Control			
Similar	2976 (1049)	3170 (1119)	3073 (1057)
Dissimilar	2826 (791)	2931 (1063)	2878 (904)
Articulatory suppression			
Similar	3127 (1178)	2836 (1117)	2981 (1160)
Dissimilar	2831 (1113)	2548 (875)	2690 (978)

On-line meta-memory: Pre-estimates of recall in any serial position.

Pre-estimates were obtained for how many words each person thought that they would recall in any serial position prior to each trial. Table 10 shows that pre-estimates of recall were lower for phonologically dissimilar than for phonologically similar words, $F(1,20) = 5.69$, $p = .027$, $MSE = .11$. Also, pre-estimates were lower when articulatory suppression was used, $F(1,20) = 123.97$, $p < .001$, $MSE = .11$, than when it was not. There was also a word-type \times articulatory suppression interaction, $F(1,20) = 10.04$, $p = .005$, $MSE = .05$. This interaction was due to lower pre-estimates for phonologically dissimilar than similar words in the articulatory suppression condition, $F(1,20) = 9.13$, $p = .007$, $MSE = .13$, but not in the control condition. Thus, people were able to predict the effect of articulatory suppression in lowering their recall, but not the effect of varying the phonological similarity of the words.

In addition to being able to predict the pattern of their performance, the second criterion was that there would be a significant predictive relationship between pre-estimates and recall in any serial position. In the control condition, both phonologically similar, $R^2 = .22$, $F(1,19) = 5.42$, $MS_{residual}$

= 0.19, and phonologically dissimilar, $R^2 = .33$, $F(1,19) = 9.28$, $MS_{\text{residual}} = 0.56$, pre-estimates showed a significant predictive relationship to the relevant performance measure at $p < .05$. In the articulatory suppression condition, neither phonologically similar nor phonologically dissimilar pre-estimates were statistically related to their counterpart performance measures.

Table 10. Mean (SD) pre-estimates of recall in any serial position for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 1 ($N = 21$).

	Pre-estimates of recall in any serial position
Control	
Similar	4.14 (.64)
Dissimilar	4.13 (.70)
Articulatory suppression	
Similar	3.51 (.82)
Dissimilar	3.18 (.75)

Despite the absence of a phonological similarity effect, persons were able to predict the effect of articulatory suppression and that phonological similarity correctly resulted in lower recall of phonologically dissimilar items in the articulatory suppression condition. Furthermore, for the control condition at least, there was a predictive relationship of pre-estimates to later recall. So, although one cannot address whether persons can or cannot predict a phonological similarity effect, because one was not present, there was some evidence from Experiment 1 that persons might be able to differentially predict working memory performance using only their perception of a change in item characteristics.

Discussion

Prior to considering the results of Experiment 1 as they impact on the three postulates being advanced in this dissertation, some discussion of the between-session

constancy, the sensitivity and bias of persons on the verification task, and the effect on the main dependent variables of verifying a sentence as either correct or muddled will be presented. The discussion will then review the results of Experiment 1 and their implications for postulates 1 to 3.

Session constancy: An issue of task reliability

In Experiment 1, the data from Sessions 1 and 2 were sufficiently similar in form that they could be aggregated together. That the data could be aggregated also illustrates that the scores on the present complex-span task are stable over a two-week period of time. Anastasi (1982) refers to this as test-retest reliability. Stability over time is important because working memory is tacitly assumed to be a stable cognitive process. Thus, a score at Time 1 should be similar at Time 2 if the general conditions of the task are the same. In summary, Experiment 1 empirically supports the assumption that working memory is a stable cognitive process and replicates the work of Tirre and Peña (1992) who also found a complex-span task to be reliable over time.

Sentence verification: Sensitivity and bias issues

Another important feature of Experiment 1 was that it was not assumed that persons were all equally able to verify the sentences correctly, nor that persons were unbiased in pressing the correct or muddled keys.

The d' measure of sensitivity (MacMillan & Creelman, 1991) was used to indicate the degree to which persons were able to correctly verify sentences. While a description of Signal Detection Theory is not appropriate here, there are some excellent applications of this procedure to tasks similar to the one described here (e.g., Grier, 1971; MacMillan & Creelman, 1991; Pastore & Scheirer, 1974). In Experiment 1 all measures of sensitivity demonstrated an extremely high ability of persons to correctly discriminate between correct and muddled sentences.

Rather than counterbalancing the key allocation to correct (C) or muddled (M) to remove any bias, Experiment 1 used the c-measure of bias (MacMillan & Creelman, 1991, to explicitly determine if persons were biased to differentially respond that a sentence was correct or muddled. The data clearly show that persons were unbiased as to which key they pressed. This lack of bias implies that persons responded on the basis of the task demands rather than other, unspecified, cognitive processes involved in choice behaviour.

In summary, Experiment 1 provides empirical evidence for assuming that the complex-span task was free of response bias and that the sentences were discriminable in terms of their veridicality.

Verifying sentences: The effect of how a sentence was verified on other dependent variables

Although the verification task was discriminable and bias free, whether a sentence was correct or muddled did alter the magnitude of the dependent variables used in Experiment 1. Measures of content were slightly lower, but measures of order and processing time were higher for correct, than for muddled items. While these results may be of interest in other contexts (e.g., studies of comprehension), without replication it is unclear what, if any, significance these results have for any of the three postulates relating to working memory.

The implications of Experiment 1 for the three postulates

Experiment 1 represented an initial foray into linking the quantitative/process and qualitative/structural models of working memory. The hope for Experiment 1 was that there would be a phonological similarity effect present that could be demonstrated in the context of a complex-span task. However, not only was there no phonological similarity effect, but persons quite unexpectedly recalled more phonologically similar than dissimilar words. Despite the

lack of a phonological similarity effect, Experiment 1 has nonetheless provided some useful insight into the operation of working memory. These insights will now be discussed separately in the contexts of Postulates 1 to 3.

Postulate 1: Demonstrating the phonological similarity effect in a complex-span task

As stated above, the most disappointing aspect of Experiment 1 was that there was no phonological similarity effect. Consequently, there is no evidence to support an assertion that the phonological similarity effect can be expected to also be apparent in complex-span tasks such as those used by Daneman and Carpenter (1980, 1983). It is unclear from Experiment 1 alone why the phonological similarity effect was not replicated. The most obvious explanation for this failure to replicate is that there is something about concurrent verification of sentences that disrupts or prevents the phonological similarity effect. This explanation is tested in Experiment 3.

It is also important to consider why persons recalled more phonologically similar than dissimilar words. One explanation may be that persons used a rule-based reconstruction strategy in recalling words. For example, if the word *cat* was remembered, then a person might be able to reconstruct the other items from knowledge that they ended in *-at*. However, such a reconstruction strategy would be unlikely to allow retrieval of information about presentation order. Consequently, it would be expected that there would be more order errors for phonologically similar than dissimilar words, as was found.

Although no phonological similarity effect was found, articulatory suppression did lower the magnitude of content measures in Experiment 1. This lowering of content measures by articulatory suppression is consistent with the proposal that articulatory suppression disrupts or prevents the

Articulatory Rehearsal Process (Baddeley, 1981, 1986, 1992a, 1992b; Baddeley, et al., 1984; Baddeley, et al., 1975).

In conclusion, the effect of articulatory suppression suggests that the Articulatory Rehearsal Process does operate in tasks based on a quantitative/process model of working memory (Postulate 1, p. 85). However, the absence of the phonological similarity effect precludes any statement about the operation of the Phonological Store in the context of a quantitative/process model of working memory.

Postulate 2: Measures of order and processing time in differentially defining working memory

Measuring order as order errors has proven to be useful in Experiment 1. The intriguing part of the data in Experiment 1 is that the loss of order information was larger for phonologically similar than dissimilar items.

The loss of order information in Experiment 1 may indicate that the more similar words are phonologically, the more confusing the order those words were presented in becomes. Also, this loss of order information occurred even in the absence of a phonological similarity effect. The effect of phonological similarity on order errors could indicate that item order is preserved in Baddeley's (1986, 1992a, 1992b) Phonological Store. Alternately, a greater number of order errors for phonologically similar than dissimilar words may have occurred because similar items were recalled via reconstruction without order information, as described above. However, even if item order were guessed, guessing accuracy would not be expected to differ across the phonological similarity conditions. That is, the proportion of order errors (out of the total number of words recalled in any serial position) ought not to be different between phonologically similar and dissimilar items. In Experiment 1, the proportion of order errors out of recall in any serial position for phonologically similar words (Mean = 0.22, SD = 0.12) was higher than the proportion for phonologically

dissimilar words (Mean = 0.15, SD = 0.15; this difference was significant, $t(20) = 3.55$, $p = .004$, one-tailed. The significantly higher proportion of order errors in the phonologically similar condition suggests that the difference in absolute order errors across word-type is not due to simply more phonologically similar words being guessed than dissimilar words. It is more likely that order errors are directly related in some way to the phonological similarity of words and not to chance.

An illustration may help clarify this point. Suppose that a person recalled 5 phonologically similar items and 4 phonologically dissimilar items. The person guessed the position of 10% of the items and was incorrect on every guess. So, the person would have obtained .5 order errors for similar and .4 order errors for dissimilar items. While the absolute number of order errors was higher for similar than for dissimilar items, the proportion of order errors out of the total recall was the same. This was not the case for Experiment 1. In Experiment 1, the proportion of order errors for phonologically similar words was higher than the proportion for phonologically dissimilar items.

In contrast, articulatory suppression had no effect on order information. The use of articulatory suppression is presumed to block the Articulatory Rehearsal Process (Baddeley, 1986, 1992a, 1992b; Baddeley, et al., 1975). This in turn suggests that perhaps preservation of item order information is separate from the operation of the Articulatory Rehearsal Process.

Processing time appears to increase when items are phonologically similar. Although increased time was spent on phonologically similar compared to dissimilar items at input, there was no predictiveness of recall by viewing and verification time.

Finally, the effect of whether a word was in a correct or muddled sentence also affected the score on other dependent variables. However, rather than presenting a treatise on the meaning of these findings here, I shall discuss them at the end of Experiment 2 when an opportunity to examine the replicability of the results has been taken.

Postulate 3: Monitoring working memory

Using the criteria that there must be a similar pattern and a relationship between on-line meta-memory and working memory performance, results from Experiment 1 tentatively confirm that persons can respond to parameter changes by altering the levels of their on-line meta-memory estimates. That is, persons were able to demonstrate some ability to predict and, by inference, monitor their working memory. However, the performance data did not show a phonological similarity effect. Thus, while Experiment 1 has shown that the potential for persons to monitor their working memory exists, nothing can be directly concluded about the relationship between meta-memory and the Phonological Store of working memory (Postulate 3, p. 88) except what was already known: that persons are able to generally predict their recall (e.g., Bruce, et al., 1982; DeVolder, et al., 1990; Hasselhorn, & Hager, 1989; Hultsch, Hertzog, Dixon, & Davidson, 1988; Leonesio & Nelson, 1990; Lovelace, 1984).

Summary

While it was disappointing not to demonstrate the phonological similarity effect in the quantitative/process model of working memory, thus supporting Postulate 1, Experiment 1 has produced some important results. First, there is some evidence of the Articulatory Rehearsal Process operating in both models of working memory. Second, measuring order and processing time in addition to content in working memory appears to be useful (Postulate 2). Third, persons demonstrated quite accurate ability in monitoring their general working memory processes (Postulate 3).

In conclusion, Experiment 1 has begun to provide an empirical base upon which to build upon. Specifically, hypotheses about order effects and meta-memory can be refined for Experiment 2.

EXPERIMENT 2: THE WORD-LENGTH EFFECT IN A COMPLEX-SPAN TASK

Like Experiment 1, Experiment 2 investigated three Postulates whose confirmation was necessary to begin to integrate two predominant ways of theorizing about working memory. Thus, Experiment 2 sought to examine if the word-length effect used to define the Articulatory Rehearsal Process of the qualitative/structural model was relevant to describing the quantitative/process model; if aspects of content, order, and processing time used in the quantitative/process model could differentiate aspects of the word-length effect more clearly; and if persons could monitor the operation of working memory sufficiently to predict the word-length and articulatory suppression effects. Experiment 2 represented a parallel study to Experiment 1. Where Experiment 1 investigated the phonological similarity effect, Experiment 2 investigated the word-length effect. Together, Experiments 1 and 2 aimed to generate data bearing on the operation of the Phonological Store and Articulatory Rehearsal Process respectively in a complex-span task.

Postulate 1: The word-length effect in a complex-span task

Experiment 2 investigated the word-length effect as an effect that Postulate 1 (p. 85) suggests might occur in both simple-span and complex-span tasks. The word-length effect of the Articulatory Rehearsal Process has been assumed to occur because, as word length increases, words take longer to sub-vocally rehearse in real time (assuming that rehearsal rate is constant). Because it takes longer to rehearse a long than a short word, more long than short words have been assumed to decay below a recall threshold before they can be rehearsed.

Consequently, fewer long than short words are likely to be recalled in serial order (Baddeley, et al., 1975). Empirical studies have found the word-length effect to be a reliable effect, having been replicated in one study of a complex-span task (La Pointe & Engle, 1990) and cross-culturally several times (e.g., Ellis & Hennelly, 1980; Hoosain & Salili, 1988; Naveh-Benjamin & Ayres, 1986). Articulatory suppression, a robust procedure that prevents the operation of the Articulatory Rehearsal Process (Baddeley, 1986, 1992a, 1992b), typically removes any word-length effect irrespective of whether words are presented visually or auditorily (Baddeley, et al., 1975; Baddeley, et al., 1984; La Pointe & Engle, 1990). It has been assumed that because the Articulatory Rehearsal Process is prevented when a person is (sub)vocally rehearsing, words are not rehearsed by the Articulatory Rehearsal Process. Consequently, recall of words in their correct order is unaffected by word length (Wickelgren, 1965). Of course, this begs the question of how any words can be recalled when the Articulatory Rehearsal Process is inoperative. Baddeley (1986) has suggested that some storage of words occurs in the Central Executive in addition to in the Phonological Store and that temporary LTM activation may be involved (A.D. Baddeley, personal communication, April 19, 1993). However, while appealing, this explanation has not yet been shown to be empirically accurate.

A potential problem with the Baddeley, et al. (1975) studies is that longer words were also words less frequently used and, by Elley's (1975) argument, less meaningful. Elley's word norms for concrete nouns were used to compare frequency between Baddeley, et al.'s long and short words. Of the eight target words in Baddeley, et al.'s first experiment that were concrete nouns, short nouns were more frequent (Mean = 5.8, SD = 3.03, N = 5) than long nouns (Mean = 8.17, SD = 0.75, N = 6). The difference in word frequency was of an extent to cause concerns about the confounding influence of frequency

and meaningfulness (Elley) on the word-length effect. However, in their third experiment, Baddeley, et al. did match short (in terms of articulatory time) and long groups of di-syllabic words for frequency. The results of their third experiment were that increasing word length, defined as time to articulate words in either the short or long groups, resulted in lower recall. It might be cautiously concluded that the word-length effect is robust enough to remain after any differences due to the frequency or meaningfulness of the target words are factored out. In the present study, however, word frequency (Elley, 1975) was equated across long and short words and across phonologically similar and dissimilar words to be on the safe side.

Postulate 1 (p. 85) implied that the word-length and articulatory suppression effects used to infer the Articulatory Rehearsal Process (Baddeley, 1986, 1992a, 1992b) would be evident in typical complex-span tasks involving an operation and recall component. This implication led to three hypotheses for Experiment 2:

1.1 There would be greater recall of one-syllable than of two-syllable words;

1.2 There would be lower recall of all words when the person was using articulatory suppression; and

1.3 There would be no word-length effect with concurrent articulatory suppression.

Together, support for Hypotheses 1.1, 1.2, and 1.3 would provide a replication of the data Baddeley, et al. (1975) originally used to infer the structure of the Articulatory Rehearsal Process. Demonstration of the presence of the Articulatory Rehearsal Process in Experiment 2 would represent an extension of Baddeley's model as it would be occurring in a complex-span task.

Finally, there was a tentative hypothesis based on the results from Experiment 1 about how the veridicality of a sentence might affect a measure of content.

1.4 Measures of content would be lower when sentences were verified as correct than when muddled.

Postulate 2: Differentiation of the word-length effect using measures of order and processing time

The rationale for investigating the utility of measures of order and processing time in Experiment 2 do not differ substantially from the rationale presented in Experiment 1 (p. 106). However, it is worth repeating the rationale as it is critical to understanding the implications for Postulate 2 (p. 88).

Regarding order information, the intention was to determine whether order errors contributed additional information about working memory in general and the word-length effect in particular. If word-length or articulatory suppression effects are found to be linked to order errors, then it could be suggested that the Articulatory Rehearsal Process is also involved in storage and rehearsal. This is perhaps unlikely as analysis of Experiment 1 showed that phonological similarity affected order errors, but not articulatory suppression, the process used to partly define the Articulatory Rehearsal Process. It is expected then, from Experiment 1, that neither the word-length nor articulatory suppression conditions will show any differences in order errors. This is because they both define the Articulatory Rehearsal Process which was apparently unaffected on the measure of order in Experiment 1. Failure of the word-length effect to demonstrate differential order errors across conditions would indicate that order information is not rehearsed nor operated upon by the Articulatory Rehearsal Process. This in turn would provide additional evidence for

the suggestion from Experiment 1 that order information is preserved in the Phonological Store in some way.

As in Experiment 1, processing time in Experiment 2 was the time spent viewing and verifying each stimulus. In Experiment 1, there was no difference in viewing and verification times between items with and without articulatory suppression. If articulatory suppression blocked rehearsal processes in Experiment 1, then the absence of change in viewing and verification times suggested that viewing time was independent of the Articulatory Rehearsal Process. So, for Experiment 2, the premise that viewing and verification time represents a change in the rehearsal time for each word was retracted. Instead, it was hypothesized that viewing and verification time was independent of the Articulatory Rehearsal Process. This hypothesis is supported by Cowan and others (Avons, et al., 1994; Caplan, et al., 1992; Caplan & Waters, 1994; Coltheart, et al., 1990; Cowan, 1992; Cowan, et al., 1992; Henry, 1991; Monsell, 1987; c.f., Baddeley & Andrade, 1994) who have suggested that the word-length effect may be an effect of output processes rather than input (rehearsal) processes.

In summary, to provide data sufficient to draw a conclusion about the utility of using order and processing time in addition to content dimensions, Experiment 2 continued investigating the effect of order errors and of viewing and verification times on the word-length and articulatory suppression effects.

Operationally, Postulate 2 implied that by using measures of content (recall in any serial position), order (order errors), and processing time (mean viewing and verification time for each target) that the Articulatory Rehearsal Process could be more exactly described. The specific hypotheses that could be derived for Postulate 2 were as follows:

2.1 There would be no effects of word-length on order errors.

2.2 There would be no effects of articulatory suppression on order errors.

Hypotheses 2.1 and 2.2 support an independence of the Articulatory Rehearsal Process from the preservation and recall of order information. Finally, based on the evidence from Experiment 1 and from other research it was hypothesized that if input and Articulatory Rehearsal Processes are separate that:

2.3 There would be no difference in viewing and verification time as a function of articulatory suppression.

Finally, as with Experiment 1, it was also argued that there might be some interaction between measures of content, order, and processing time with whether a sentence is verified as correct or muddled. For the data from Experiment 1 to be replicated, correct verifications would produce lower scores on the content measures than muddled verifications, and correct verifications would produce higher scores on measures of order and processing than muddled verifications.

Postulate 3: Prediction of the word-length effect through monitoring of working memory

The final aim of Experiment 2 was to examine whether persons were able to predict their own performance as a function both of word-length and of articulatory suppression. In fact, data from Experiment 1 indicated that persons were able to predict that articulatory suppression would lower recall.

Operationally, Postulate 3 (p. 88) implied that content measures of on-line meta-memory (Kausler, 1991) would mirror content measures of memory performance in both pattern and relationship. Again, while it could not be hypothesized

whether persons would be able to predict the word-length effect, it was predicted in Experiment 2 that:

3.1 Persons would be able to predict the effect of articulatory suppression on the pattern of recall in any serial position.

Summary

Experiment 2 aimed to take a group of results used to define part of the Phonological Loop, the Articulatory Rehearsal Process, and to examine whether those results also occurred in a complex-span task. Experiment 2 also examined whether dimensions of content, order, and processing time were useful in differentiating aspects of the word-length effect. Finally, Experiment 2 examined whether persons were able to monitor their task performance in terms of recall. These three aims empirically tested Postulates 1 to 3 respectively, and paralleled the aims of Experiment 1.

Method

Participants

Fourteen female participants aged 19 - 42 (Mean = 25.3 years; SD = 6.2) and 7 male participants aged 19 - 28 (Mean = 22.9; SD = 3.6) completed Experiment 2. The participants in Experiment 2 were the same participants as for Experiment 1.

Design

Experiment 2 was a three-way factorial design, with all factors varied within subjects. The first factor was word-type (1-syllable, 2-syllable), the second factor was articulatory suppression (absent or present), and the third factor was sentence type (correct or muddled). Experiment 2 was conducted in two sessions separated by an interval of two

weeks. The dependent variables were as detailed in the General Method section (p. 99).

Procedure

The procedure for Experiment 2 followed that described in the General Method section (p. 95) and in Experiment 1 (p. 109).

Results

Reliability

As with Experiment 1, stability over time in Experiment 2 was calculated using the Pearson product-moment correlation coefficient. Table 11 shows the correlation coefficients for each dependent variable as a function of articulatory suppression (control, present) and word-type (1-syllable, 2-syllable) conditions. For all the dependent variables of recall in the correct serial position, recall in any serial position, verification time, and pre-estimates, test-retest reliability over the two week interval was moderate to high in all conditions. From these data, it can be concluded that measures of working memory remain stable over short time periods, even with small sample sizes and only three trials per session.

The data for both sessions in Experiment 2 were also compared across all of the dependent variables using a MANOVA. None of the dependent variables showed any statistical differences for the main effect of session across the two-week interval. Nor were there any interactions between session and any of the other dependent variables. As none of the dependent variables showed an effect of practice over the two-week period, the data were again aggregated. All further reports and analyses of data are thus based on aggregated data.

Table 11. A summary of the two week test-retest reliability of the complex-span task for 1 and 2-syllable words with and without articulatory suppression in Experiment 2. Test-retest reliabilities were moderate to high in all conditions and were calculated using the Pearson product-moment coefficient.

	Recall in the correct serial position ^{*1}	Recall in any serial position ^{*1}	Verification times ^{*2}	Pre-estimates ^{*2}
Control				
1-syllable	.59**	.51*	.80***	.73***
2-syllable	.52*	.55**	.86***	.72***
Articulatory suppression				
1-syllable	.53*	.53*	.78***	.71***
2-syllable	.82***	.61**	.93***	.80***

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

*1 N. = 21

*2 N = 18

Finally, in Experiment 2, an analysis of persons' sensitivity and bias in responding to the sentence verification task was again performed. In Experiment 2, participants had very high sensitivity and responded without any bias (Appendix 2; Table A.8). That is, participants were able to correctly detect whether a sentence was correct or muddled and to respond without bias.

Content Measures: Recall in the correct serial position and recall in any serial position.

As with Experiment 1, Experiment 2 used two measures of content: recall in the correct serial position and recall in any serial position.

Recall in the correct serial position

Hypothesis 1.1 predicted that there would be greater recall of 1 than of 2-syllable words, a hypothesis unsubstantiated by the data. It was also expected that articulatory suppression would lower recall overall (Hypothesis 1.2), which did occur, $F(1,20) = 42.17$, $p < .001$, $MSE = .34$ (Figure 6; Table 12). There was no word-type by articulatory suppression interaction, thus failing to confirm Hypothesis

1.3. The failure to find an interaction may have been due to the absence of a word-length effect in the control condition.

There was an effect of verification condition such that words verified as correct were better recalled than words verified as muddled, $F(1,20) = 44.94$, $p < .001$, $MSE = .19$, the opposite of what was predicted by Hypothesis 1.4. Finally, there was a word-type \times articulatory suppression \times verification interaction, $F(1,20) = 6.52$, $p = .019$, $MSE = .07$. This interaction occurred because there was a word-type \times articulatory suppression interaction in the muddled verification condition, $F(1,20) = 6.21$, $p = .022$, $MSE = .17$, but not in the correct verification condition (Figure 6; Table 12).

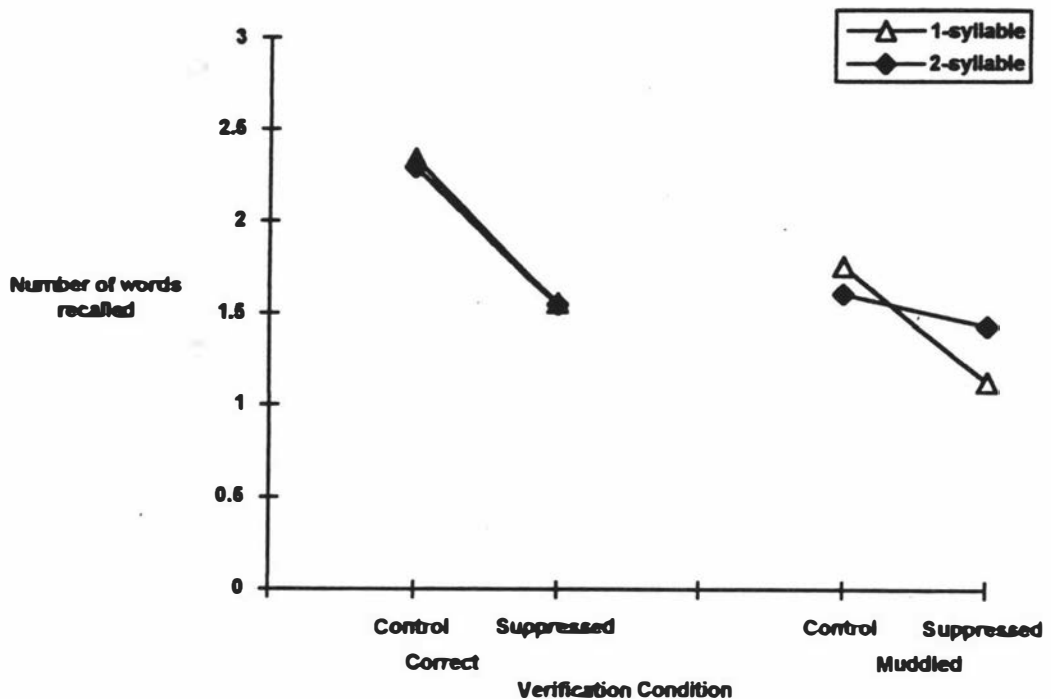


Figure 6. Recall in the correct serial position by word-length, articulatory suppression, and verification conditions for Experiment 2.

Recall in any serial position

When examining recall in any serial position, again there was no word-length effect. Again articulatory suppression lowered recall overall, $F(1,20) = 43.91$, $p < .001$, $MSE = .37$. There

was a word-length x articulatory suppression interaction, $F(1,20) = 9.56$, $p = .006$, $MSE = .08$. This interaction occurred because there was slightly higher recall of 1-syllable than of 2-syllable words in the control condition and slightly lower recall of 1-syllable than of 2-syllable words in the articulatory suppression condition.

There was also higher recall of words verified as correct than as muddled, $F(1,20) = 25.04$, $p < .001$, $MSE = .15$, and a word-length x verification interaction, $F(1,20) = 8.59$, $p = .008$, $MSE = .04$. The higher recall of words in sentences verified as correct than muddled contradicts Hypothesis 1.4. The interaction occurred because there was a word-length effect in the muddled verification condition, $F(1,20) = 12.42$, $p = .002$, $MSE = .04$, but not in the correct verification condition. There was also an interaction between word-type and articulatory suppression in the muddled verification condition, $F(1,20) = 14.78$, $p = .001$, $MSE = .08$, which occurred because there was a word-length effect with, but not without, articulatory suppression. The presence of this word-type x articulatory suppression interaction in the muddled but not correct verification condition produced a word-type x articulatory suppression x verification condition interaction, $F(1,20) = 5.40$, $p = .031$, $MSE = .09$ (Table 12).

Table 12. Mean (SD) recall of 1-syllable or 2-syllable words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions in Experiment 2. Note that recall is out of three for either correct or muddled verification conditions and out of 6 for the total verification and pre-estimates columns ($N = 21$).

	Recall in the correct serial position (SD)			Recall in any serial position (SD)		
	Correct	Muddled	Total	Correct	Muddled	Total
Control						
1-syllable	2.34 (.48)	1.75 (.74)	4.10 (1.19)	2.49 (.42)	2.17 (.50)	4.66 (.86)
2-syllable	2.29 (.52)	1.60 (.74)	3.90 (1.11)	2.44 (.47)	2.09 (.58)	4.53 (.90)
Articulatory suppression						
1-syllable	1.56 (.70)	1.13 (.61)	2.68 (1.24)	1.80 (.54)	1.34 (.59)	3.13 (1.05)
2-syllable	1.55 (.78)	1.43 (.71)	2.98 (1.42)	1.81 (.61)	1.75 (.54)	3.56 (1.08)

Analysis of recall in any serial position produced a perhaps confusing array of results. In summary, these data showed that the only hypothesis supported was that articulatory suppression lowers recall in any serial position, in the same manner as occurred for recall in the correct serial position. Typically, persons recalled more words from sentences that were semantically correct than muddled. It is unclear what, if any, implications the other interactions have. To begin to draw conclusions from these interactions at this point would at best be premature.

Summary of the data relating to Postulate 1

Experiment 2 failed to find a word-length effect. In Experiment 2, there was no evidence that word length has a detrimental effect on a person recalling the content or content and order of the word. There was again evidence that articulatory suppression lowered recall of words overall. Because there was no word-length effect in the first instance, there could not be a removal of a word-length effect by articulatory suppression. Finally, in contrast to Experiment 1, Experiment 2 showed that verifying a word as correct is likely to produce better recall than does verifying a word as muddled. Interpretation of these apparently contradictory findings will be discussed below.

Order Measures: Order errors

Order errors were again calculated as the number of words recalled in any serial position minus the number of words recalled in their correct serial position (Table 13).

As predicted by Hypothesis 2.1, there was no difference in order errors as a function of word-type, $F(1,20) = 2.26$, $p = .149$, $MSE = .06$. Neither was there any difference in the number of order errors as a function of articulatory suppression, $F(1,20) = 3.88$, $p = .063$, $MSE = .05$ (Hypothesis 2.2). Thus, these data supported the suggestion that the

Articulatory Rehearsal Process is independent of the process by which order information is maintained.

Contrary to the results obtained in Experiment 1, in Experiment 2 there were a larger number of order errors in sentences verified as muddled than correct, $F(1,20) = 13.21$, $p = .002$, $MSE = .08$. Finally, there was a verification \times articulatory suppression interaction, $F(1,20) = 8.53$, $p = .008$, $MSE = .10$, produced by the tendency of order errors to decrease when sentences were verified as correct and increase when sentences were verified as muddled.

Table 13. Mean (SD) order errors for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 2. Order errors are out of 3 for the correct and muddled conditions, and out of 6 overall ($N = 21$).

	Mean order errors (SD)		
	Correct	Muddled	Total
Control			
1-syllable	0.15 (0.21)	0.41 (0.39)	0.56 (0.52)
2-syllable	0.15 (0.25)	0.48 (0.42)	0.63 (0.54)
Articulatory suppression			
1-syllable	0.24 (0.36)	0.21 (0.23)	0.45 (0.41)
2-syllable	0.26 (0.39)	0.32 (0.34)	0.58 (0.61)

Processing time: Verification and viewing time

As with Experiment 1, Experiment 2 used the time to verify a sentence as either correct or muddled as a measure of processing time. Data from Experiment 2 showed that persons spent more time viewing 1-syllable than 2-syllable stimuli, $F(1,20) = 74.53$, $p < .001$, $MSE = 79,964$ (Table 14). It is not readily apparent why this should have been so.

If, as asserted earlier, the Articulatory Rehearsal Process is independent of the processes operating at input in the present task, then it is reasonable to have expected that articulatory suppression would have had no effect on viewing times of the stimuli (Hypothesis 2.3). As predicted, data from Experiment 2 showed that articulatory suppression had no

effect on the time stimuli were viewed. Furthermore, regression equations failed to show any predictive relationship between viewing time and recall in the correct serial position, further strengthening a case that input processes are unrelated to the Articulatory Rehearsal process.

The only other significant effect was that verification time was faster when sentences were verified as correct than when verified as muddled, $F(1,20) = 6.94$, $p = .016$, $MSE = 168,477$ (Table 14).

In summary, the only finding of relevance to the present discussion is that the time taken at initial processing of stimuli (input processes) did not appear to be affected by articulatory suppression. This result is consistent with the interpretation that articulatory suppression operates on post-input processes and that the input processes are at best minimally involved with the Articulatory Rehearsal Process.

Table 14. Mean (SD) verification times (ms) for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 2 (N = 21).

	Verification times (ms; SD)		
	Correct	Muddled	Mean
Control			
1-syllable	2882 (898)	3092 (1072)	2987 (952)
2-syllable	2540 (895)	2673 (1082)	2606 (974)
Articulatory suppression			
1-syllable	2937 (1196)	3074 (1327)	3006 (1242)
2-syllable	2539 (1113)	2727 (1229)	2633 (1160)

On-line meta-memory: Pre-estimates of Recall in Any Serial Position

Pre-estimates were obtained for how many words each person thought that they would recall in any order prior to each trial (Table 15).

There was only one hypothesis relating to on-line working memory monitoring: that persons would be able to predict the

effect of articulatory suppression on their pattern of recall in any serial position (Hypothesis 3.1). The data from Experiment 2 provided support for this hypothesis through lower pre-estimates for articulatory suppression than no articulatory suppression, $F(1,20) = 52.43$, $p < .001$, $MSE = .29$ (Table 15). Because there were no effects of word-length, it was unsurprising that the pre-estimates of recall did not predict a word-length effect.

Table 15. Mean (SD) pre-estimates of recall in any serial position for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 2 (N = 21).

	Pre-estimates
Control	
1-syllable	4.14 (0.80)
2-syllable	4.11 (0.82)
Articulatory suppression	
1-syllable	3.33 (.090)
2-syllable	3.22 (0.87)

The data also showed that in each condition there was a significant relationship between pre-estimates and recall. Specifically, in the control condition pre-estimates for both 1-syllable, $R^2 = .54$, $F(1,19) = 21.83$, $MS_{residual} = 0.36$, and 2-syllable, $R^2 = .53$, $F(1,19) = 21.05$, $MS_{residual} = 0.40$, words were significantly related to performance at $p < .05$. Also, in the articulatory suppression conditions, 1-syllable, $R^2 = .67$, $F(1,19) = 38.29$, $MS_{residual} = 0.38$, and 2-syllable, $R^2 = .49$, $F(1,19) = 18.04$, $MS_{residual} = 0.63$, words were significantly related to performance at $p < .05$.

In summary, the data again showed that persons were able to produce on-line meta-memory estimates that reflected both the pattern of performance and were demonstrably related to performance.

Discussion

Before discussing the implications of the data from Experiment 2 for the three postulates being used to integrate definitions of working memory, the stability of measures of working memory over time, issues of sensitivity and bias, and the effect of how a sentence was verified will be discussed. As much of the discussion will be no more than a repetition of points made in Experiment 1, the discussion of these three issues will be brief.

Session constancy: An issue of task reliability

In Experiment 2, the data from Sessions 1 and 2 were again sufficiently similar in form that they could be aggregated together. That the data could be aggregated also illustrates that the scores on the present complex-span task are stable over a two-week period of time. In summary, Experiment 2 empirically supports the assumption that working memory is a stable cognitive process and replicates the results of Experiment 1.

Sentence verification: Sensitivity and bias issues

Another important feature of Experiment 2 was that, like Experiment 1, it was not assumed that persons were all equally able to verify the sentences correctly, nor that persons were necessarily unbiased in pressing the correct or muddled keys.

The measure of sensitivity (d' ; MacMillan & Creelman, 1991, in Experiment 2 demonstrated that persons had an extremely high ability to correctly discriminate between a correct and a muddled sentence. Like Experiment 1, Experiment 2 used the c-measure of bias (MacMillan & Creelman, 1991) to determine if persons were biased to differentially respond that a sentence was correct or muddled. The data also showed that persons were practically unbiased as to which key they pressed. This low bias implies that, in Experiment 2, persons again responded on the basis of the demands of the task

rather than other cognitive processes involved in choice behaviour.

In summary, Experiment 2 replicated Experiment 1 in providing empirical evidence that the present task was free of response bias and that the sentences were discriminable in terms of their veridicality.

Verifying sentences: The effect of how a sentence was verified on other dependent variables

Although the verification task was discriminable and bias free, whether a sentence was correct or muddled did alter the magnitude of the dependent variables used in Experiment 2. However, in Experiment 2, the direction of these changes was opposite in every instance to Experiment 1. Measures of content were slightly higher, but measures of order and processing were lower for correct than muddled items. Because these results contradict those from Experiment 1, it is very unclear what significance these results have for the operation of working memory. This issue will be discussed further in the summary section below (p. 146).

The implications of the results of Experiment 2 for the three general postulates about working memory

Experiment 2, like Experiment 1, provides empirical support for Postulates 2 and 3. However, there is no basis to confirm that the effects used to define the qualitative/structural model of working memory were evident in the present complex-span task, a task based on the quantitative/process model of working memory (Postulate 1).

Postulate 1: Demonstrating the word-length effect in a complex-span task

In order to support the postulate that the operation of the quantitative/process model of working memory reflects some or all of the dimensions of the qualitative/structural model of working memory, Experiment 2 needed to exhibit effects definitive of the Articulatory Rehearsal Process. Word length

and articulatory suppression effects are critical to inferring operation of the Articulatory Rehearsal Process. Experiment 2 was disappointing in that no word-length effects were exhibited. Thus, no statement in support of Postulate 1 (p. 85) can be made.

The second prediction necessary to linking the present complex-span task with the Articulatory Rehearsal Process was that articulatory suppression, which is hypothesized to block the Articulatory Rehearsal Process, would lower recall in the present complex-span task (Baddeley, et al., 1984; Baddeley, et al., 1975; Vallar & Baddeley, 1984). As expected, articulatory suppression did lower overall recall of target items. However, while this finding is necessary in inferring the Articulatory Rehearsal Process, it is not sufficient on its own to do so.

The third expectation in Experiment 2 was that word-length effects would disappear when articulatory suppression was used by persons, because articulatory suppression prevents operation of the Articulatory Rehearsal Process. The presupposition of this expectation is that a word-length effect is present before articulatory suppression. Because no word-length effects were present before articulatory suppression, the absence of word-length effects after articulatory suppression cannot be claimed as evidence of anything about articulatory suppression and its role in removing word-length effects. Hence, no comment about the operation of the Articulatory Rehearsal Process in the present complex-span task can be made.

Thus, the way articulatory suppression was operating in the present complex-span task is indeterminate. Without a word-length effect in the first place, articulatory suppression may simply create a divided attention task which interferes with or slows the operation of working memory, hence lowering recall.

From the present data then, it is impossible to determine whether the Articulatory Rehearsal Process is operating in the present complex-span task. This also means that one cannot conclude from Experiment 2 that it is possible to demonstrate the qualitative effect of the Articulatory Rehearsal Process in the present quantitative model. Thus Experiment 2 has failed to provide support for Postulate 1. However, Experiment 2 has not disconfirmed Postulate 1 either. So the question is open as to why the word-length effect was absent in Experiment 2.

There are at least two possible reasons why Experiment 2 failed to demonstrate word-length effects. The first is that the presence of the sentence verification task in some way precluded operation of the Articulatory Rehearsal Process and that the effect of articulatory suppression was in fact due to task interference.

The second reason for the absence of word-length effects may have been that the discriminability of the stimuli was insufficient: that is, that 1 and 2-syllable words are just not different enough in (articulatory) length to produce a word-length effect. The reason that this hypothesis was not investigated next was that there was insufficient evidence to conclude that the concurrent operation had no impact on the word-length effect. The hypothesis that sentence verification had a bearing on the absence of a word-length effect seemed likely when placed in the context of the results from Experiment 1, in which no phonological similarity effect was present with a concurrent task. It seemed, at the time, that the evidence obtained in Experiments 1 and 2 required direct investigation of the sentence verification task.

Postulate 2: Measures of order and processing time in differentially defining working memory

In Experiment 2, neither the length of the words nor the presence or not of articulatory suppression had any effect on order information. This is an important result. It was found

in Experiment 1 that order information was affected by phonological similarity but not by articulatory suppression. The suggestion to be taken from this result is that order information is related in some way to the Phonological Store, but not to the Articulatory Rehearsal Process. Thus, if order information is unrelated to the Articulatory Rehearsal Process, then in Experiment 2, neither manipulations across word length nor articulatory suppression conditions ought to have any effect on order information. This is exactly what was found in Experiment 2. There is thus some basis for being more certain that order information is not part of the Articulatory Rehearsal Process and hence advancing the hypothesis that the Articulatory Rehearsal Process may operate solely on item content.

The only effect on processing time of any of the manipulations in Experiment 2 was that more time was spent viewing and verifying the 1-syllable words than the 2-syllable words. It is unclear what implications this result has as it cannot be placed in a context of any statements about the Articulatory Rehearsal Process. If articulatory suppression affirms the presence of the Articulatory Rehearsal Process, then it might be concluded that processing time at input and the Articulatory Rehearsal Process are separate processes (e.g., Cowan, 1992; Cowan, et al., 1992).

Finally, whether a sentence was correct or muddled again showed some effect on other dependent variables. In Experiment 2, correct sentences had higher scores on measures of content than muddled sentences. Correct sentences had lower scores on order and processing measures than muddled sentences. The results in Experiment 2 contradict the results from Experiment 1. Consequently, there seems to be no coherent pattern to the results. The only immediate conclusion that can be drawn is that it is critical to balance the presentation correct and muddled sentences within each trial to avoid confounding the results in any way.

Postulate 3: Monitoring working memory

Finally, in Experiment 2, persons were able to predict the pattern of their performance. Also, in addition to predicting the pattern of their performance, persons' on-line meta-memory appeared to be related to their actual performance in some regular way. This result provides further support for the possibility of monitoring one's own working memory performance. However, as a word-length effect was absent in Experiment 2, it remains to be determined if persons can actually predict that effect, and hence the operation of the Articulatory Rehearsal Process in working memory.

SUMMARY OF EXPERIMENTS 1 AND 2

Experiments 1 and 2 were designed to provide a first step toward integrating the quantitative/process and qualitative/structural models of working memory. To integrate these models it was first necessary to demonstrate operation of the Phonological Loop (comprised of the Phonological Store and the Articulatory Rehearsal Process) in operationalisations of both models. In Experiment 1, the presence of the phonological similarity effect without but not with articulatory suppression was the criterion for demonstrating the Phonological Store. In Experiment 2, the presence of the word-length effect without but not with articulatory suppression was the criterion for demonstrating the operation of the Articulatory Rehearsal Process. However, neither Experiment 1 nor 2 were able to replicate the phonological similarity and word-length effects, respectively, in a complex-span task. So, at this point, there is no direct evidence that the Phonological Loop operates in the quantitative/process model as it does in the qualitative/structural model of working memory.

The phonological similarity and word-length effects have been repeatedly demonstrated in simple-span tasks. However, Experiments 1 and 2 of the present set of studies failed to

demonstrate the phonological similarity and word-length effects in a complex-span task, a task that could be described as a simple-span task plus a concurrent operation. Given that the effects the present studies sought to replicate have been described as robust, the question is raised as to what effect the presence of a concurrent operation might be having on the working of the Phonological Loop. The logical way to examine the effect of a concurrent operation is to repeat Experiments 1 and 2 without the concurrent operation, that is, try to replicate the effects used to infer the Phonological Loop in a *simple-span* task.

Both Experiments 1 and 2 demonstrated that it is useful to use theoretical features of content, order, and processing time in investigating the operation of working memory (Postulate 2). The main implication of this result is that working memory cannot be simply described as a storage area nor solely in terms of its capacity (c.f., Cantor & Engle, 1993; Engle, et al., 1990; Kyllonen & Christal, 1990; La Pointe & Engle, 1990; Light & Anderson, 1985; Turner & Engle, 1989; Whitney, Ritchie, & Clark, 1991; Wingfield, Stine, Lahar, & Aberdeen, 1988). From Experiments 1 and 2, working memory appears to be multi-dimensional in nature. Experiments 1 and 2 also suggest that working memory involves information about what is active and what order that activation occurs in. There is thus support for Postulate 2 that:

2. The factors of order and processing time of process/quantitative models will be useful in elaborating upon the qualitative/structural model of working memory, in addition to a measure of content.

Finally, Experiments 1 and 2 have extended research in which persons' were shown to predict their recall on other than working memory tasks (e.g., Bruce, et al., 1982; DeVolder, et al., 1990; Perlmutter, 1978; Rebok & Balcerak, 1989). Experiments 1 and 2 has extended this research by

demonstrating that persons are able to *differentially* predict the pattern of their working memory performance and that these predictions are related to their performance. What remains to be determined is whether this relationship between meta-memory and working memory remains when the phonological similarity and word-length effects are replicated.

In conclusion, Experiments 1 and 2 found direct evidence for Postulates 2 and 3 only. This implies that dimensions of content, order, and processing time are useful in describing working memory and that persons can monitor the products of their working memory. What remains to be determined is whether the Phonological Loop occurs in both qualitative/structural and quantitative/process models of working memory.

CHAPTER 8: EXPERIMENTS 3 & 4: EXAMINATION OF QUANTITATIVE/PROCESS, QUALITATIVE/STRUCTURAL, AND SUBJECTIVE/OBJECTIVE FACTORS IN A SIMPLE-SPAN TASK

Experiments 3 and 4 were conducted to answer a single important question: Was the absence of phonological similarity and word-length effects in the first two experiments due to the concurrent processing aspect of the task, that is, sentence verification? A criticism of Experiments 1 and 2 was that, while the 'storage' aspect of working memory was being studied, this storage aspect may have been confounded by what Just and Carpenter (1992) describe as the processing aspect of the Reading Span task, namely the verification of sentences. Data in La Pointe and Engle's (1990) first two experiments showed smaller magnitude word-length effects for complex-span than simple-span tasks. Thus is it not unreasonable to assert that the concurrent sentence operation interferes with effects used to define the Phonological Loop.

However, there are also data showing that the difference between using simple or complex-span tasks is not that large or theoretically relevant. For example, Verhaegen, Marcoen, and Goossens (1993) analyzed 40 studies of memory span examining the differences between complex-span and simple-span tasks in the context of age. Verhaegen, et al. found that both complex-span and simple-span measures of working memory displayed similar effect sizes in discriminating between age groups. This result suggests that the supposed advantage in predicting other cognitive performance from a complex-span task, in preference to a simple-span task, is not as empirically sound as it is theoretically appealing. In the present dissertation, the relevance of Verhaegen, et al.'s meta-analysis is that perhaps the concurrent operation of sentence verification was not responsible for the absence of phonological similarity and word-length effects in Experiments 1 and 2.

Thus, given one set of research suggesting that the concurrent operation is critical to working memory measurement and another minimizing its impact, it is empirically sensible to repeat Experiments 1 and 2 but without the concurrent verification task. Finally, Experiments 3 and 4 also provided an opportunity to replicate the results of Experiments 1 and 2 with respect to Postulates 2 and 3.

EXPERIMENT 3: THE PHONOLOGICAL SIMILARITY EFFECT IN A SIMPLE-SPAN TASK

As explained above, the primary rationale for conducting Experiment 3 was to investigate whether the phonological similarity effect may have been obscured in Experiment 1 because of the presence of a concurrent operation (Figure 3, p. 87). The evidence for this explanation is based on theory (Just & Carpenter, 1992) and empirical work in which a concurrent operation has reduced the word-length effect (La Pointe & Engle, 1990).

However, it is also important to note that not only was there no phonological similarity effect in Experiment 1, but that there was greater recall of phonologically similar than of phonologically dissimilar items. As was discussed earlier, one likely explanation for this was that persons used a rule-based strategy to guess the words, with a consequent absence of order information. It could be argued that it would be sensible to try to reduce the operation of this rule-based strategy. One way to do so might be to reduce the time persons had for recall, for example. However, at the time of conducting Experiment 3, varying more than one parameter might have also prevented a conclusion of whether the absence of a phonological similarity effect in Experiment 1 was obscured by concurrent sentence verification or the other parameter(s) that were varied. As the impact of concurrent operations upon assessment of working memory is both theoretically and empirically important, it was decided to

change only the concurrent operation parameter in Experiment 3.

In Experiment 1 it was stated that Postulate 1 implied that the phonological similarity and articulatory suppression effects used to infer the Phonological Loop structure (Baddeley, 1986, 1992a, 1992b) would be evident in complex-span tasks involving an operation and storage component. In Experiment 1, this assertion was unsupported. In Experiment 3, it was expected that the phonological similarity and articulatory suppression effects used to define the Phonological Store would occur in a simple-span task where they did not in a complex-span task. Thus, the hypotheses relating to the phonological similarity effect in Experiment 3 were unchanged from those in Experiment 1:

1.1 There would be greater recall of phonologically dissimilar than similar items;

1.2 There would be lower recall of all items when the person was using articulatory suppression; and that

1.3 There would be no phonological similarity effect with concurrent articulatory suppression.

The second result that Experiment 3 sought to replicate was the larger effect on order errors of phonologically similar than phonologically dissimilar words. As was shown in Experiment 1, this effect was unlikely to be explained solely by guessing alone. The proportion of order errors for phonologically similar words was significantly larger, indicating that phonological similarity had an effect at input or maintenance of words beyond any effect of guessing order at output. So, in Experiment 3, it was predicted that:

2.1 phonologically similar items would produce more order errors than phonologically dissimilar items.

Also, in both Experiments 1 and 2, articulatory suppression had no effect on a measure of order. So, in Experiment 3, it was expected that:

2.2 articulatory suppression would not alter the number of order errors.

The measure of processing time was less interesting in Experiments 1 and 2 of the present dissertation. When processing time was defined as the time spent viewing and verifying an item, it did not appear to have any real impact on any other variables. Experiment 3 provided an opportunity to replicate the lack of impact of processing time upon recall but without the additional sentence verification time.

Finally, Experiment 3 aimed to replicate the ability of persons to monitor their working memory performance. From Experiments 1 and 2, it was expected to find accurate monitoring of working memory performance patterns, and a relationship between meta-memory and performance. In Experiment 3 the scene was set to examine whether this accuracy could be replicated when a phonological similarity effect was present. Thus, provided that performance data showed a phonological similarity effect, it was hypothesized in Experiment 3 that:

3.1 Persons would be able to predict the patterns of the phonological similarity effect on their recall in any serial position.

Method

Participants

Twelve female participants aged 23 - 48 (Mean = 30.40 years; SD = 7.65) and eight male participants aged 19 - 41 (Mean = 31.00; SD = 8.40) completed Experiment 3. Due to a computer error, viewing time and pre-estimate data were lost for two

participants reducing the sample size from 20 to 18 for those measures.

Design

Experiment 3 was a two-way factorial design, with all factors varied within subjects. The first factor was word-type (phonologically similar or phonologically dissimilar) and the second factor was articulatory suppression (absent or present). Experiment 3 was conducted in a single session. The dependent variables were as detailed in the General Method section (p. 99).

Procedure

The procedure for Experiment 3 was identical to the procedure for Experiment 1 with one exception. When the stimulus was presented, there was no verification task to complete. So, to view each succeeding stimulus, the participant pressed the space-bar on the keyboard. Otherwise the procedure has been described fully in the General Method section (p. 95).

Results

Results of Experiment 3 will be presented in terms of measures of content, order, processing time, and of meta-memory.

Content Measures: Recall in the correct serial position, recall in any serial position

The results of Experiment 3 and its investigation of Postulate 1 will now be presented in terms of recall in the correct serial position and recall in any serial position (Table 16). The results again failed to demonstrate a phonological similarity effect, but did show a lowering of measures of content with concurrent articulatory suppression.

Table 16. Mean (SD) recall of phonologically similar or phonologically dissimilar words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions in Experiment 3 (N = 20).

	Recall	
	Recall in the correct serial position	Recall in any serial position
Control		
Similar	3.92 (1.18)	5.22 (0.74)
Dissimilar	3.97 (1.30)	4.70 (0.65)
Articulatory suppression		
Similar	2.85 (1.38)	4.02 (0.84)
Dissimilar	2.75 (1.43)	3.45 (1.10)

Recall in the correct serial position

In order to demonstrate a phonological similarity effect, Hypothesis 1.1 predicted that phonologically dissimilar words would be better recalled than phonologically similar words, an expectation that was not confirmed by the data over the control and articulatory suppression conditions, $F(1,19) = .02$, $p = .890$, $MSE = .64$. Items that were phonologically similar and phonologically dissimilar were recalled equally well. There was lower recall when articulatory suppression was used by persons, $F(1,19) = 29.49$, $p < .001$, $MSE = .88$, supporting Hypothesis 1.2. As there was no phonological similarity effect without articulatory suppression, the absence of an interaction (as implied by Hypothesis 1.3) between phonological similarity and articulatory suppression was unsurprising. In short, Experiment 3 was unable to replicate the phonological similarity effect, thus again failing to provide support for Postulate 1 (Table 16; Figure 7).

Recall in any serial position

As was found in Experiment 1, data in Experiment 3 again showed increased recall for phonologically similar over phonologically dissimilar words, $F(1,19) = 71.77$, $p < .001$, $MSE = .58$ (Hypothesis 1.1). As predicted by Hypothesis 1.2, articulatory suppression lowered the overall level of recall, $F(1,19) = 24.22$, $p < .001$, $MSE = .24$. Finally, there was no

phonological similarity x articulatory suppression interaction. Again, recall in any serial position failed to demonstrate a phonological similarity effect in the present simple-span task (Table 16; Figure 7).

To summarise, contrary to predictions based on Baddeley's (1986, 1992a, 1992b) qualitative/structural model of working memory, phonological similarity effects in recall continued to be absent in the simple-span task presented in Experiment 3. The main implication of this result is that *it is unlikely that the presence of the concurrent task precluded the phonological similarity effect in Experiment 1.*

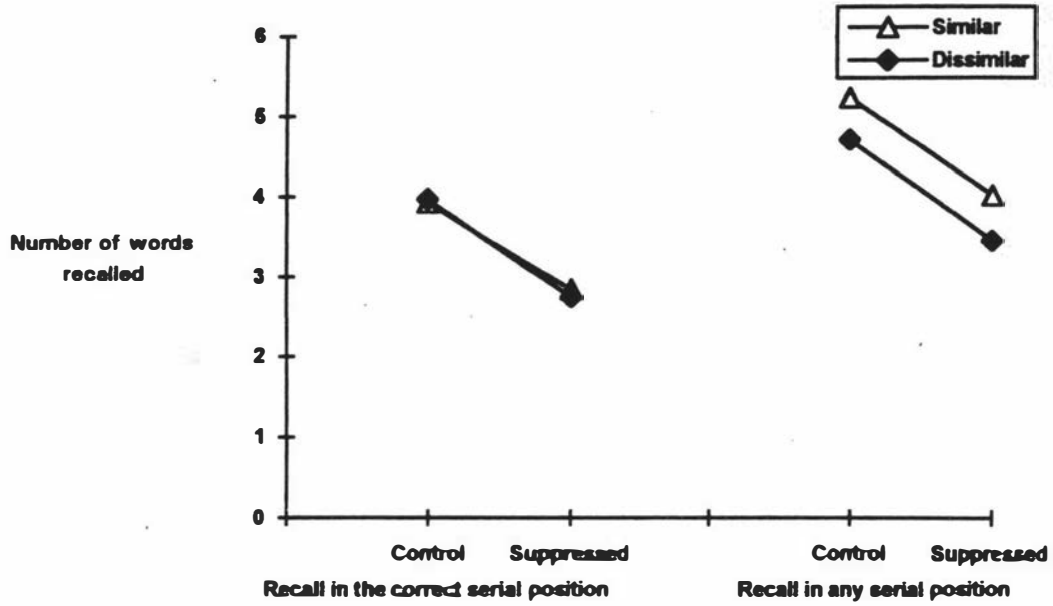


Figure 7. Recall in the correct serial position and recall in any serial position by phonological similarity and articulatory suppression conditions in Experiment 3.

It was interesting in Experiment 3 to find that recall in the correct serial position was equivalent across the phonological similarity conditions, whereas phonologically similar words were better recalled than phonologically dissimilar words for recall in any serial position. This difference in effects replicates research using free-recall

methods that showed that when only content was required for recall, phonologically similar items were better recalled than phonologically dissimilar items (Wickelgren, 1965). The implication of this result is that increased recall of phonologically similar items over phonologically dissimilar items is a function of content. As will be demonstrated below, lower recall of phonologically similar than phonologically dissimilar items may be produced by a loss of order information.

Order Measures: Order errors

The data in Experiment 3 showed, as expected, that there were more order errors for phonologically similar than phonologically dissimilar items (Hypothesis 2.1), $F(1,19) = 12.51$, $p = .002$, $MSE = .43$. As predicted by Hypothesis 2.2, there was no effect of articulatory suppression (Table 17).

Table 17. Mean (SD) order errors for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 3 (N = 20).

	Order errors
Control	
Similar	1.30 (0.92)
Dissimilar	0.73 (0.82)
Articulatory suppression	
Similar	1.17 (0.92)
Dissimilar	0.70 (0.65)

Because recall levels were different in one measure of content, it could be proposed that there were more order errors because more items were recalled, but that the proportion of order errors would be equivalent. Anticipating this criticism, Experiment 3 found that the proportion of order errors (order errors divided by the total words recalled in any serial position) was greater for phonologically similar (Mean = .28, SD = .18) than for phonologically dissimilar (Mean = .20, SD = .21) words, $t(19) = 2.14$, $p = .023$, one-tailed, across articulatory

suppression condition. This increased proportion of order errors for phonologically similar over phonologically dissimilar words indicates again that order errors are differentially affected by the phonological similarity of the stimuli.

In summary, Experiment 3 supports an assertion that order information is separate from the effects of articulatory suppression, and by implication, separate from the Articulatory Rehearsal Process (Table 17). The implication of this finding is that increasing phonological similarity reduces a person's ability to maintain order information, a result that replicates that of Experiment 1.

Processing time: Viewing time

In Experiment 3, there were no significant differences in mean viewing time as a function of phonological similarity. Unexpectedly, articulatory suppression lowered the time spent viewing the target words, $F(1,17) = 9.51$, $p = .007$, $MSE = 538\ 759$ (Table 18). This lowered viewing time is perplexing because intuitively one would expect that engaging in articulatory suppression, an additional processing task, would demand that more, not less, time be spent rehearsing and reading words. An explanation provided by one participant was that he chose to go faster when he was using articulatory suppression because he believed that he would forget fewer words this way. This lowered time spent processing items with articulatory suppression in Experiment 3 is also at odds with the null effects of articulatory suppression found in Experiments 1 and 2. Therefore, at this point, this anomalous effect will be treated as an aberration pending further empirical investigation in subsequent experiments.

Finally, in Experiment 3 the standard deviations were typically quite large. The minimum viewing times across all conditions ranged from 623 ms to 969 ms and the maximum from 6,229 ms to 7,687 ms. However, it was not practical to remove

any especially long viewing times as there would then have been too few cases to analyze.

Table 18. Mean (SD) viewing times (ms) for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 3 (N = 18).

	Viewing times (ms)
Control	
Similar	3165 (2006)
Dissimilar	2829 (2100)
Articulatory suppression	
Similar	2450 (1717)
Dissimilar	2478 (1867)

On-line meta-memory: Pre-estimates of Recall in Any Serial Position

Hypothesis 3.1 predicted that persons would be able to predict the patterns of the phonological similarity effect on their recall. However, as the performance data from Experiment 3 failed to replicate the phonological similarity effect, one could not expect to support Hypothesis 3.1. The data did show that Experiment 3 pre-estimates were higher for the control than the articulatory suppression condition, $F(1,17) = 13.21$, $p = .002$, $MSE = .62$. There were no other main or interaction effects (Table 19). So, only the pattern of articulatory suppression was predicted in Experiment 3. There was no prediction of the higher recall of phonologically similar over phonologically dissimilar words.

Table 19. Mean (SD) pre-estimates of recall in any serial position for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions in Experiment 3 (N = 18).

	Pre-estimates
Control	
Similar	4.20 (1.00)
Dissimilar	4.37 (1.02)
Articulatory suppression	
Similar	3.57 (1.11)
Dissimilar	3.65 (1.05)

There was a statistically significant relationship at $p < .05$ found between pre-estimates and performance measures for phonologically similar and for phonologically dissimilar words in the control condition, $R^2 = .28$, $F(1,15) = 5.77$, $MS_{\text{residual}} = 0.23$, and $R^2 = .43$, $F(1,15) = 11.29$, $MS_{\text{residual}} = 0.25$, respectively. There was also a statistically significant relationship at $p < .05$ between pre-estimates and performance measures for phonologically similar and for phonologically dissimilar words in the articulatory suppression condition, $R^2 = .50$, $F(1,15) = 14.80$, $MS_{\text{residual}} = 0.36$, and $R^2 = .65$, $F(1,15) = 27.93$, $MS_{\text{residual}} = 0.50$, respectively. These data replicated previous results (Experiments 1 and 2) in which on-line meta-memory had shown a relationship to general working memory performance.

Discussion

The principal conclusion from Experiment 3 is that the presence of a concurrent sentence verification task is probably not the reason that phonological similarity effects were absent in Experiment 1. The implication of this conclusion is that (a) this research program must establish the conditions for replicating the phonological similarity effect, and (b) that an absence of a phonological similarity effect is not likely due to the presence of a concurrent operation.

The absence of a phonological similarity effect in Experiment 3 is in some ways both disappointing and welcome. It is disappointing because the question of why phonological similarity effects were absent in Experiments 1 and 3 remains an open question. It is welcome because if the presence of a concurrent operation precluded the phonological similarity effect, there would be little basis to claim that qualitative aspects of working memory could be demonstrated in complex-span tasks such as those used by the quantitative/process model of working memory. If presence of phonological

similarity was precluded by a concurrent operation, then clearly both models of working memory would appear to be considering different constructs and processes. However, despite an absence of a phonological similarity effect in both Experiments 1 and 3, the results from Experiment 3 bear scrutiny in respect of the three postulates of this dissertation.

Postulate 1: Demonstrating the phonological similarity effect in a simple-span task

Experiment 3 failed to demonstrate a phonological similarity effect. That is, where phonologically similar items were expected to be less well recalled than phonologically dissimilar items, in Experiment 3 they were recalled equally well in the correct serial position. In any serial position, more phonologically similar than phonologically dissimilar items were recalled. The question is raised as to why Experiments 1 and 3 have failed to replicate the phonological similarity effect.

In thinking about why these present studies have not replicated the phonological similarity effect, the author began with two hypotheses: (a) that the phonological similarity effect was obscured because persons had too much time to recall items; or, (b) that the phonological similarity effect only occurs for items taken from a very small pool of stimuli and presented more than once.

It was discussed in Experiment 1 that persons may have been reconstructing recall of items from a rule-based system for phonologically similar items only. That is, if a person recalled one word, they could then try adding consonant prefixes onto that word's stem to see if they could guess or recognize the word. For example, if the first word was *cat*, a person could generate words such as *bat*, *hat*, *pat*, and recognize that *pat* was the second target item. The prediction for such a strategy would be that persons might recall more

phonologically similar than phonologically dissimilar items by reconstruction but that when the order those items occurred in was considered, recall would be equivalent (Wickelgren, 1965). These predictions are borne out by the results of Experiment 3. The inference to be made from these results is that perhaps reconstruction could be limited by limiting the time persons have at recall. Examining the research done with item recall does suggest that it is usual to allow only between 10 to 20 seconds to recall items (e.g., Baddeley, 1966a, 1968; Baddeley, et al., 1984; Baddeley, et al., 1975). Thus, in future research it may be necessary to limit recall time rather than allow an unlimited time to recall items, in order to reduce item reconstruction.

The second hypothesis for the absence of a phonological similarity effect is derived from an examination of research in which the phonological similarity effect has been demonstrated. This examination found that, in most studies of phonological similarity effects, the stimulus pool was of about 10 items which were sampled from with replacement (especially Baddeley, 1966a, 1968; Baddeley, et al., 1984; Baddeley, et al., 1975; Conrad & Hull, 1964; Estes, 1973; Healy, 1974). LaPointe and Engle (1990) in their fourth and fifth experiments showed that articulatory suppression removed the effects of word-length for a fixed stimulus pool but not for a pool from which items were sampled without replacement. Taken together, the results from previous research do suggest that the nature of the stimulus pool is worth examination.

Both these issues of reconstruction and pool-type deserve more investigation and will be examined more fully in the next chapter. However, to recapitulate, data from Experiment 3 are interpreted as disconfirming the hypothesis that a concurrent verification task is responsible for the absence of a phonological similarity effect.

Postulate 2: Measures of order and processing time in differentially defining working memory

As with Experiment 1, Experiment 3 demonstrated that measures of order and processing time, in addition to a measure of content, are useful in further defining dimensions of working memory. Measures of content showed that articulatory suppression lowers the amount of material in working memory. This is consistent with both Articulatory Rehearsal Process and interference interpretations of the role of articulatory suppression. Without specifically knowing whether articulatory suppression also blocks the phonological similarity effect, it is impossible to determine how articulatory suppression lowers measures of content from the present data.

Experiment 3 replicated the results in Experiment 1 wherein more order information was lost for phonologically similar than for phonologically dissimilar items. This result occurred even in the absence of a phonological similarity effect. This is an important result as evidence is now beginning to converge that the phonological similarity effect might be produced by a loss of order, not content, information. The implication of the phonological similarity effect being an effect of successive order errors is that the Phonological Store of the qualitative/structural model of working memory cannot account for order information (Burgess & Hitch, 1992) and, from this inference, should do so. Experiment 3 highlights that to fully describe working memory, it appears at least prudent to include some measure of how effectively item order is preserved (Postulate 2, p. 88).

Finally, in Experiment 3, the expectation that input and working memory processes would remain separate was partially supported by the data. It would be logical to suggest that if articulatory suppression either blocks the Articulatory Rehearsal Process or interfered in some other way with

acquisition of information, that the time taken to input items would increase. Paradoxically, the time taken to input items decreased with articulatory suppression. While one participant explained that he went faster with articulatory suppression to avoid a loss of information, what exactly occurred is at this stage unknown. To validate or disconfirm any theories based on this effect of processing time on Experiment 3 there would need to be at least a clearer replication of this effect.

In summary, Experiment 3 continued to support the utility in describing working memory in terms of three dimensions of content, order, and processing time. These dimensions provide a means to dissociate and describe the processes of working memory more clearly at both operational and theoretical levels. Particularly, the differential loss of order information as a function of phonological similarity appears to be assuming greater theoretical significance in considering working memory.

Postulate 3: Monitoring working memory

Finally, Experiment 3 again provided support for the assertion that working memory output can be monitored. In Experiment 3, persons predicted their pattern of recall in response to using articulatory suppression. Furthermore, the on-line pre-estimates were predictive of actual performance.

However, because there was no phonological similarity effect in performance data in Experiment 3, it remains unclear whether persons will be able to predict a phonological similarity effect when it is present. Thus, when the phonological similarity effect is replicated, it will be of interest to test whether persons predict that effect or not. To summarize, Experiment 3 adds a third confirmation of persons' ability to monitor their general working memory performance accurately.

EXPERIMENT 4: THE WORD-LENGTH EFFECT IN A SIMPLE-SPAN TASK

The main purpose of Experiment 4 was to examine whether the absence of the word-length effect in Experiment 2 was due to the concurrent sentence verification operation (Figure 3, p. 87). In Experiment 2, it was expected that the word-length and articulatory suppression effects used to infer the Articulatory Rehearsal Process (Baddeley, 1986, 1992a, 1992b) would be evident in typical complex-span tasks involving an operation and recall component. However, there was no word-length effect in the data from Experiment 2, hence there was no basis to conclude that the Articulatory Rehearsal Process operated in a complex-span task. If the presence of a concurrent operation in Experiment 2 was what prevented replication of a word-length effect, then in the simple-span task of Experiment 4 there was reason to expect replication of a word-length effect. Therefore, the following hypotheses were proposed:

- 1.1 There would be greater recall of 1-syllable than 2-syllable items;
- 1.2 There would be lower recall of all items when the person were using articulatory suppression; and that
- 1.3 There would be no word-length effect with concurrent articulatory suppression.

Experiment 4 also continued to use additional measures of order and processing time (Postulate 2, p. 88). In Experiments 1, 2 and 3, the assessment of item order preservation produced data that differentiated between phonological similarity and word-length effects. Increased phonological similarity of items produced increased order errors in Experiments 1 and 3. However, increased word-length and articulatory suppression did not produce any change in order errors in Experiment 2. If this differential increase in order errors, as a function of phonological similarity but

not word-length or articulatory suppression, indicates the dissociability of the Phonological Store from the Articulatory Rehearsal Process respectively, then for Experiment 4 it might be expected that:

2.1 Word-length and articulatory suppression manipulations will not alter the number of order errors.

As in Experiments 1 and 2, Experiment 3 again showed that the measure of processing time was not effective in delineating phonological similarity or word-length effects. The data from Experiment 3 only partially supported the hypothesis that the input processes (as assessed by viewing time) were separate from the processes supposedly operating in the Articulatory Rehearsal Process (articulatory suppression). That is, processing time decreased when articulatory suppression was used relative to when it was not. From an inspection of the data, it is probable that this effect of articulatory suppression on viewing times was due to two reasons: one person reported deliberately going faster with articulatory suppression so as to reduce the total time suppressing; and, some persons spent an inordinate amount of time on some trials in the control condition as shown by the large standard deviations in Table 18. With these explanations in mind, in Experiment 4 the hypothesis was re-proposed that:

2.2 There would be no difference in viewing times as a function of word-type, indicating the separateness of the Articulatory Rehearsal Process from input processes, and

2.3 There would be no differences between viewing times for items when articulatory suppression was used than when it was not.

From Experiments 1, 2 and 3, data continued to support Postulate 3 which implies that content measures of on-line meta-memory will mirror content measures of memory

performance in pattern and be related to those performance measures. Thus it was hypothesized in Experiment 4 that, provided that a word-length effect was present, that:

3.1 Persons would be able to predict the pattern of a word-length effect on their recall in any serial position.

In addition, it was hoped that Experiment 4 would show a word-length effect. Until a word-length effect is replicated, one cannot address whether persons can monitor such an effect.

Method

Participants

Nineteen of the twenty participants in Experiment 3 completed Experiment 4 also. Eleven female participants aged 23 to 48 (Mean = 29.78 years; SD = 7.84) and eight male participants aged 19 - 41 (Mean = 31.00; SD = 8.40) completed Experiment 4. Due to a computer error, viewing time and pre-estimate data were lost for two and one participants respectively, reducing the sample size from 19 to 17 and 18 accordingly for those measures.

Design

Experiment 4 was a two-way factorial design, with all factors varied within subjects. The first factor was word-type (1-syllable, 2-syllable) and the second factor was articulatory suppression (absent or present). The dependent variables were as detailed in the General Method section (p. 99).

Procedure

The target words were identical to those in Experiment 2, except that the words were presented on their own and not at

the end of a sentence. Otherwise the procedure was identical to Experiment 2 and as described in the General Method section (p. 95).

Results

Content Measures: Recall in the correct and recall in any serial position

As with Experiments 1 to 3, the results for measures of content will be presented separately for recall in the correct serial position and recall in any serial position (Table 20).

Table 20. Mean (SD) recall of 1-syllable or 2-syllable words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions in Experiment 4 (N = 19).

	Measures of content	
	Recall in the correct serial position	Recall in any serial position
Control		
1-syllable	4.84 (1.43)	5.28 (0.72)
2-syllable	4.54 (1.26)	5.11 (0.76)
Articulatory suppression		
1-syllable	3.11 (1.22)	4.09 (0.90)
2-syllable	3.33 (1.45)	4.26 (0.95)

Recall in the correct serial position

From predictions based on a qualitative/structural model, it was expected that 2-syllable words would be less well recalled than 1-syllable words (Hypothesis 1.1), an expectation that was not confirmed by the data, $F(1,18) = 0.11$, $p = .749$, $MSE = .22$. Hypothesis 1.2 predicted that articulatory suppression would lower recall of items overall, which the data showed in Experiment 4, $F(1,18) = 32.99$, $p < .001$, $MSE = 1.25$ (Table 20; Figure 8). As with Experiment 2, there were no word-length effects for articulatory suppression to prevent, and so Hypothesis 1.3 could not be addressed.

Recall in any serial position

There was no difference in recall in any serial position between 1-syllable and 2-syllable words (failing to confirm Hypothesis 1.1), $F(1,18) = 0$, $p = 1.000$, $MSE = .18$. Articulatory suppression lowered recall overall, $F(1,18) = 28.47$, $p < .001$, $MSE = .69$ (Hypothesis 1.2). Finally, there was no interaction between word-type and articulatory suppression conditions, $F(1,18) = 1.73$, $p = .205$, $MSE = .34$ (Hypothesis 1.3).

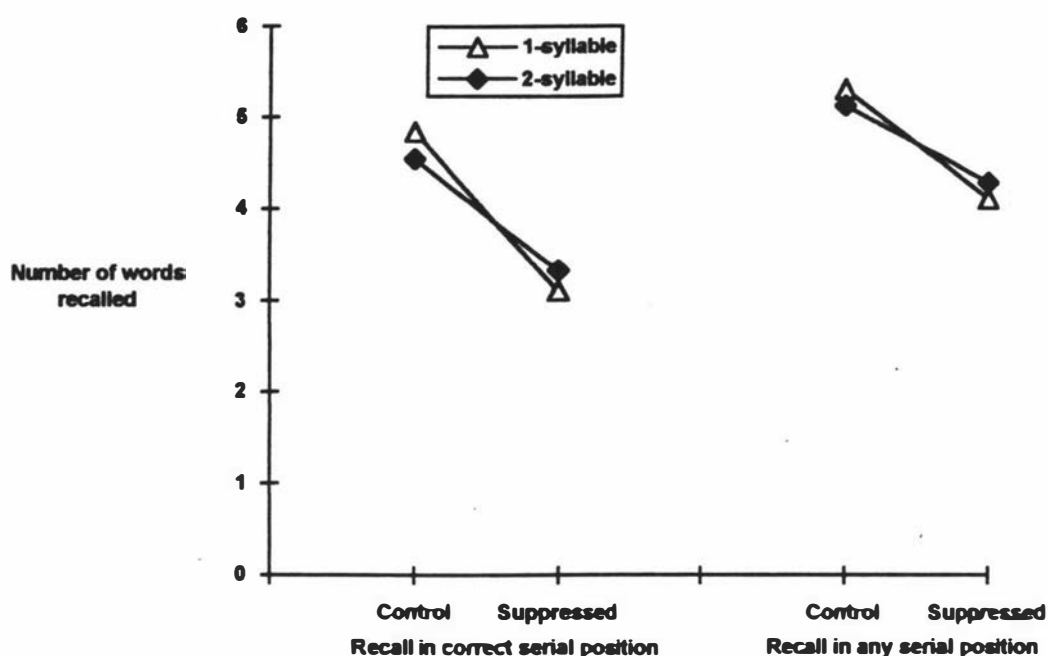


Figure 8. Recall in the correct serial position and recall in any serial position by word-length and articulatory suppression conditions in Experiment 4.

Summary

For the measures of content in Experiment 4, essentially the same results as Experiment 2 were found. There were no word-length effects to indicate the operation of the Articulatory Rehearsal Process in the present simple-span task, just as they were absent in the complex-span task. The main point of Experiment 4 was to examine if the absence of the word-length effect in Experiment 2 was due to the presence of a concurrent operation in addition to the span task. The data

from Experiment 4 show that the absence of the word-length effect in the complex-span task was not because of the presence of a concurrent operation.

Order Measures: Order errors

Overall, there was no difference in order errors between 1-syllable and 2-syllable words. Unexpectedly, more order errors were made with articulatory suppression than in the control condition, $F(1,18) = 12.45$, $p < .01$ (Table 21), a finding counter to Experiments 1, 2, and 3 in which articulatory suppression had no effect on order errors. It is unclear from the data why this should have been so, and this result awaits replication when the conditions for producing a word-length effect in recall are established. Thus, Hypothesis 2.1 that word-length and articulatory suppression would have no effect on order errors was only partially supported by the data from Experiment 3.

Table 21. Mean (SD) order errors for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 4 ($N = 19$).

	Order errors (SD)
Control	
1-syllable	0.44 (0.86)
2-syllable	0.56 (0.84)
Articulatory suppression	
1-syllable	0.98 (0.73)
2-syllable	0.93 (0.91)

Of primary importance to the present research, however, is that the number of order errors did not change due to word-length differences. While this may seem unsurprising because there were also no recall differences due to word-length, remember that order errors did increase as a function of phonological similarity despite the absence of a phonological similarity effect in Experiments 1 and 3. However, it remains to be determined whether order errors change when a word-length effect is present.

Processing time: Viewing time

As predicted by Hypothesis 2.2 there were no differences in mean viewing time as a function of word-type and, as predicted by Hypothesis 2.3, there were no differences between viewing times for items when articulatory suppression was used relative to when it was not (Table 22). Taken together, these results do support the assertion that the processes operating at input, and assessed by viewing time, do not impinge on the operation of the Articulatory Rehearsal Process. However, this conclusion must be tempered by the fact that there was only an effect of articulatory suppression in Experiment 4. Thus, it remains to be examined whether, with both word-length and articulatory suppression effects, input processes remain separate from maintenance and output processes.

Table 22. Mean (SD) viewing times (ms) for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 4 (N = 17).

	Viewing time
Control	
1-syllable	2513 (2217)
2-syllable	2766 (2585)
Articulatory suppression	
1-syllable	2493 (2501)
2-syllable	2691 (2870)

A second feature of the viewing times in Experiment 4 was that again the standard deviations were quite large. The minimum viewing times across all conditions ranged from 556 ms to 641 ms and the maximum from 8,211 ms to 11,294 ms. As with data in Experiment 3, it was again not practical to remove any especially long viewing times as there would have then been too few cases to analyse. These large deviations beg the question of what participants are doing during these long viewing times, an issue that will be returned to in Experiment 6.

On-line meta-memory: Pre-estimates of recall in any serial position

Because there was no word-length effect in the present Experiment 4, it could not be expected that pre-estimates would indicate one, as they did not. Thus, Hypothesis 3.1 could not be directly addressed. As can be seen from Table 23, persons did predict the detrimental effect of articulatory suppression on their recall, $F(1,17) = 3.08$, $p = .097$, $MSE = 2.05$, although only at a statistical trend level. So, although less convincingly than in the previous three experiments, persons in Experiment 4 were able to predict the pattern or effect of articulatory suppression on their recall.

Table 23. Mean (SD) pre-estimates of recall in any serial position for 1-syllable and 2-syllable words for both the control and articulatory suppression conditions in Experiment 4 (N = 18).

	Pre-estimates
Control	
1-syllable	4.30 (1.00)
2-syllable	4.37 (1.02)
Articulatory suppression	
1-syllable	3.57 (1.11)
2-syllable	3.65 (1.05)

When the relationship between pre-estimates and recall in any serial position was examined, in all but one instance in which confidence was only a statistical trend level, pre-estimates were all significantly related to recall. Specifically, there was a statistically significant relationship at $p < .05$ found between pre-estimates and performance measures for 1-syllable words in the control condition, $R^2 = .28$, $F(1,15) = 5.86$, $MS_{residual} = 0.44$. There was also a significant relationship at $p < .05$ between pre-estimates and performance measures for 1-syllable and for 2-syllable words in the articulatory suppression condition, $R^2 = .50$, $F(1,15) = 15.01$, $MS_{residual} = 0.46$, and $R^2 = .31$, $F(1,15) = 6.88$, $MS_{residual} = 0.64$, respectively. So, although there was no word-length effect in Experiment 4, hence

precluding direct examination of the relationship between pre-estimates and recall, there were nonetheless some interesting results. Again, the data showed that both the pattern and relationship of pre-estimates to recall was such that it could be concluded that persons were monitoring their working memory, at least in a general manner.

Discussion

The main result of Experiment 4 is that, even without a concurrent operation, such as sentence verification, the word-length effect was not replicated. Thus, it also seems unlikely that a word-length effect was absent in Experiment 2 solely because the experimental task involved a concurrent operation. While this absence of a word-length effect is perplexing, it is also encouraging as there clearly remains a basis to try to integrate both qualitative/structural and quantitative/process models of working memory. That is, because the absence of a word-length effect was not demonstrated as due to a concurrent operation, there is no basis to conclude that a word-length effect might not occur in both simple-span and complex-span tasks.

Notwithstanding the failure of Experiment 4 to find a word-length effect, consideration of the results of Experiment 4 in terms of Postulates 1 to 3 of this dissertation achieves three functions. First, reasons for a failure to replicate a word-length effect are explored. Second, the utility of a multi-dimensional assessment of working memory is replicated. And third, persons are again shown to be able to monitor general working memory performance.

Postulate 1: Demonstrating the word-length effect in a simple-span task

Experiment 4 generally failed to provide any support for the first postulate that predicts that a word-length effect ought to be present in both simple-span and complex-span tasks. However, while the word-length effect was absent in the

complex-span task in Experiment 2, it was also absent in the simple-span task in Experiment 4. From this, the data in Experiment 4 suggest that the dimension encompassed by including a concurrent operation in addition to a span component is not responsible for the absence of the word-length effect.

This conclusion forces one to consider other explanations for why a word-length effect has not been replicated. One possible explanation for the lack of word-length effects might be that the task was participant-paced. This explanation is supported by the large variability in viewing times. Participant-pacing allowed participants to choose the amount of time to spend rehearsing a word, and presumably forming the best memory possible. Some research has used self-paced presentation of a complex-span task and produced data with a word-length effect (La Pointe & Engle, 1990). However, given that the typical mode of presentation in which word-length effects are demonstrated is an experimenter-paced mode (e.g., Baddeley, et al., 1984; Baddeley, et al., 1975), it is clearly important to check whether pacing of stimuli exerts an influence on the word-length effect and hence the Articulatory Rehearsal Process.

A second reason for the absence of a word-length effect might be a lack of experimental power. One effect that a lack of power may have had was that the difference between 1-syllable and 2-syllable words might have been insufficient to produce word-length effects. This is the issue of treatment effects. Unfortunately, an analysis of power could not be performed for Experiment 4 because the mean sum of squares for the effect of word-length was smaller than the mean square error term. The formula for calculating power (Keppel, 1982) is only effective when the mean square of the effect is greater than the mean square error.

In summary, while it is impossible to state from the data in Experiment 4 why the word-length effect has been absent, it

is possible that participant-pacing and a loss of statistical power may be contributing factors.

Postulate 2: Measures of order and processing time in differentially defining working memory

Experiment 4 has continued to add information about Postulate 2 (p. 88) which asserts that measures of order and processing, in addition to content, are useful in conceptualizing working memory dimensions more effectively.

Experiment 4 showed that articulatory suppression increased order errors. This result was unexpected as it had been suggested from Experiments 1 to 3 that articulatory suppression, a process blocking the rehearsal of items in the Articulatory Rehearsal Process, acted only on item content. From Experiment 4, it would seem that order information may not be as separate from the Articulatory Rehearsal Process as first proposed, although the data are clearly in need of replication when a word-length effect is present.

Also, Experiment 4 showed that articulatory suppression had no effect on viewing time at input, and by implication the Articulatory Rehearsal Process appears unrelated to processing time. This is an important result and confirms the results from Experiments 1 and 2 of the present series. This result also casts doubt on the theoretical validity of the surprising result in Experiment 3 where viewing times increased with articulatory suppression. Thus, Experiment 4 continues to provide converging evidence that input and rehearsal processes may be separate from output processes (Cowan, 1992; Cowan, et al., 1992).

Postulate 3: Monitoring working memory

Finally, Experiment 4 continues to show that on-line meta-memory for operation of working memory in general is moderately accurate. Also, persons appear able to predict the deleterious effect of articulatory suppression on recall.

These results will be considered below, along with results from Experiment 3.

SUMMARY

Experiments 3 and 4 were conducted to examine whether the absence of either the phonological similarity or the word-length effect in Experiments 1 and 2, using a complex-span task, was due to the interfering effect of the sentence verification procedure. The general conclusion is that the concurrent operation is not likely to be the reason that those effects were not replicated in Experiments 1 and 2. This conclusion means that the question remains open as to whether qualitative effects (phonological similarity and word-length) can co-exist in the definition of working memory which is described as having both a span and operation component. Thus, while Experiments 1 to 4 do not allow acceptance of the assertion that aspects of the qualitative/structural model also occur in the quantitative/process model of working memory, the results do not falsify Postulate 1 either.

Utilizing Dimensions of Order and Processing Time in Addition to Content Information

In Experiment 3, order information was again shown to be less well recalled with phonologically similar than with phonologically dissimilar items. In contrast, order information in Experiment 4 was independent of the length of the words. The implication of these results is that the phonological similarity and word-length effects might operate differently in respect of order information, and thus suggest that perhaps the Phonological Store and Articulatory Rehearsal Process can be dissociated using a measure of order information.

In Experiments 3 and 4, the impact of increased viewing times on any of the other measures of working memory appeared to be

minimal. Viewing time as a measure of processing time is increasingly seeming more vague, if not irrelevant, to the concept of working memory. Could it be that, provided a person operates within a set of parameters (a minimum and maximum viewing time), that working memory can operate equally effectively? If this is so, then what is the role of processing time in the operation of working memory? Perhaps more fundamentally, what is processing time? The data gathered thus far do not allow answers to these questions. However, these questions will be borne in mind as the present research proceeds. In conclusion, Experiments 3 and 4 continue to provide evidence for the utility of taking measures of different working memory dimensions and thus support Postulate 2.

Meta-memory for Working Memory

Finally, Experiments 3 and 4 have again demonstrated that persons are able to accurately predict their own working memory performance at a general level. In addition, persons were also sufficiently sensitive to predict the effect of articulatory suppression on that recall. Although it was disappointing to not be able to directly examine whether persons were monitoring specific working memory processes because both the phonological similarity and word-length effects were unreplicated in Experiments 3 and 4, these same experiments did replicate the effects found in Experiments 1 and 2 regarding on-line meta-memory. Thus, in considering Experiments 1 to 4, one is able to advance the claim that persons do indeed appear to possess meta-memory for their general working memory and that persons possess meta-memory for the deleterious effect of articulatory suppression on their working memory.

It now remains to be examined whether this ability to monitor one's working memory includes the ability to monitor qualitative processes such as the phonological similarity and

word-length effects. To do this, one must of course have data in which performance measures also demonstrate phonological similarity and word-length effects. However, with the results from Experiments 1 to 4, there exists a set of replicated results which can be built upon in succeeding research.

At this point one can speculate that if the effect of articulatory suppression on content information is assumed to indicate the operation of the Articulatory Rehearsal Process, then it does seem that changes in the operation of the Articulatory Rehearsal Process might be able to be actively monitored. To examine this assertion further, there does need to be a clear replication of the word-length effect and a concurrent demonstration of accurate monitoring of that effect, as the Articulatory Rehearsal Process cannot be inferred from the effect of articulatory suppression alone.

Conclusions

In conclusion, Experiments 3 and 4 have helped fill in some of the gaps in the knowledge of how the qualitative/structural and quantitative/process models of working memory are related. The next stage in the present research is to investigate under what conditions both the phonological similarity and word-length effects can be replicated. Answering this question is a prerequisite to addressing more fully the question of whether the two concepts of working memory are able to be integrated.

CHAPTER 9: REPLICATING THE PHONOLOGICAL SIMILARITY AND WORD-LENGTH EFFECTS IN A SIMPLE-SPAN TASK:

The primary purpose of the experiments reported in the present chapter is to *identify the parameters within which phonological similarity and word-length effects are replicable* (Figure 3, p.87). So far, Experiments 1 to 4 have found results supportive of Postulate 2, concerning the impact of order information and processing time, and of Postulate 3, concerning persons' meta-memory of their working memory. In addition to supporting Postulates 2 and 3, it is critical to an integration of working memory models that phonological similarity and word-length effects are demonstrated in both simple-span and complex-span tasks (Postulate 1). However, before a comparison across task-type can be performed, phonological similarity and word-length effects must be replicated in a simple-span task.

Chapter 9 has three main sections. The first section presents Experiment 5 in which two stimulus sampling methods are compared. The phonological similarity effect is found to be more robust when stimuli are repeated in a different order for each trial than with new stimuli for each trial. This suggests that the phonological similarity effect results from order confusion.

The second section of this chapter presents Experiment 6 in which the discriminability of word-lengths and the pacing of presentation is varied. One and three-syllable length words produced a word-length effect in Experiment 6 where 1- and 2-syllable words did not in Experiments 2 and 4. Recall in Experiment 6 was unaffected by whether the stimuli were participant-paced or experimenter-paced.

The third section of this chapter is a summary of the major findings of Experiments 5 and 6. It is concluded that replication of the phonological similarity and word-length effects in Experiments 1 to 4 was precluded because of a need

to refine the presentation parameters of the stimuli, rather than due to more serious theoretical issues.

EXPERIMENT 5: A COMPARISON OF TWO STIMULUS SAMPLING METHODS ON THE PHONOLOGICAL SIMILARITY EFFECT

The major purpose of Experiment 5 was to examine the nature of the phonological similarity effect and the parameters necessary for its replication. Until the phonological similarity effect is replicated, Postulate 1 (p. 85) cannot be directly addressed. The second focus of Experiment 5 was to extend the previous research (Experiments 1 to 4) into the utility of measures order and processing time, in addition to content. Specifically, it was hoped to investigate measures of order and processing time when a phonological similarity effect was present (Postulate 2, p. 88). The third focus of Experiment 5 was to further examine the efficacy of using on-line-metamemory estimates of working memory output. As with previous experiments in this research program, Experiment 5 compared the patterns and relationships of working memory performance with on-line meta-memory estimates (Postulate 3, p. 88).

The phonological similarity effect

The phonological similarity effect is said to occur when stimuli (e.g., letters, words) that are phonologically similar are more poorly recalled than stimuli that are not phonologically similar (Chapter 2). This phonological similarity effect disappears with articulatory suppression for visually, but not auditorily, presented words.

However, the phonological similarity effect, and hence evidence for the Phonological Store, have thus far been absent in both a complex-span (Experiment 1), and a simple-span task (Experiment 3). The first purpose of Experiment 5 was to investigate whether the absence of the phonological similarity effect in complex and simple-span tasks was solely

because of differences in methodology or because of more serious theoretical differences between qualitative/structural and quantitative/process models of working memory. If methodological differences obscured the phonological similarity effect in both the complex and simple-span tasks, then specifying under what conditions the phonological similarity effect will be apparent in the present simple-span task would allow a synthesis of the two paradigms. That is, if Experiment 5 replicates the phonological similarity effect, then the path is open for examining the integration of the two working memory models by testing if the phonological similarity effect occurs in both simple and complex-span tasks.

To help form hypotheses about why the phonological similarity has not been replicated thus far, a meta-analysis of Experiments 1 and 3 was conducted. (The data from Experiments 1 and 3 were combined using SPSS.PC's JOIN command.) A new variable GROUP was formed to distinguish between the simple and complex variants used in Experiments 1 and 3 respectively. Joining the data was justifiable because the sole difference between Experiments 1 and 3 was the change from sentences plus words to words alone.

The combined data for Experiments 1 and 3 showed no difference in levels of recall in the correct or in any serial position between experiments (the GROUP variable). That is, whether or not the task was a complex-span or simple-span task made no difference to recall in the correct serial position. Overall, phonologically similar words were no better recalled in the correct serial position than phonologically dissimilar words. In contrast, for recall in any serial position, the combined data from Experiments 1 and 3 showed that phonologically similar words were better recalled than phonologically dissimilar words, $F(1,39) = 67.00$, $p < .001$, $MSE = .28$. The main point of these results is that, when order plus content information was required,

there was no effect of phonological similarity. In contrast, when only the content of the items was required, phonologically similar items were better recalled than phonologically dissimilar items.

Finally, articulatory suppression lowered recall in the correct serial position, $F(1,39) = 80.67$, $p < .001$, $MSE = .62$ and in any serial position, $F(1,39) = 144.36$, $p < .001$, $MSE = .40$. There were no other interactions nor main effects.

This meta-analysis of Experiments 1 and 3 suggests two reasons why the phonological similarity effect was not replicated. First, reconstruction of words may have confounded the effect of phonological similarity. The evidence for this suggestion was that when recall did not require any order information, phonologically similar words were better recalled than phonologically dissimilar words. When order was required, there was no difference in recall between phonologically similar and phonologically dissimilar words. Thus, the greater recall of phonologically similar than of phonologically dissimilar words can be explained by the indefinite recall period during which items can be reconstructed using the base sound for the word and trying different consonants for 'fit' (Wickelgren, 1965). But, although this suggestion explains greater recall of phonologically similar than phonologically dissimilar words, it does not explain why the phonological similarity effect (lower recall of phonologically similar than phonologically dissimilar words) was not replicated. As will be discussed below, it is possible that the phonological similarity effect was absent because of the nature of the stimulus sampling procedures in Experiments 1 and 3.

Reconstruction of words

Reconstruction of words, facilitated by an unlimited recall period, was hypothesized as the reason why more phonologically similar than phonologically dissimilar items were recalled. It was noted earlier that the enhanced recall

of phonologically similar over dissimilar words may have been because recall for phonologically similar words was assisted by a strategy in which participants reconstructed words in the recall phase. Thus, if the target words in order were *dog*, *log*, *cog*, and *bog*, participants need only recall the stem of the word *-og* and try various permutations on that stem to obtain at least some of the moderately frequent original stimuli. So, recall could easily reflect reconstruction of what are quite frequent words. However, while reconstruction can produce the words themselves, reconstruction as a strategy cannot reproduce the order that the words were presented in. Thus, it could be inferred that the enhanced recall of phonologically similar over dissimilar words would be absent when recalling the order of the words was an additional requirement of the recall task. This hypothesis was supported by the meta-analysis of Experiments 1 and 3. Thus, in Experiment 5, it was hypothesized that reconstruction of words might be prevented by limiting the time participants have to recall the stimuli (Baddeley, et al., 1975; Baddeley, et al., 1984).

Phonological similarity: Loss of content or of order

As noted, preventing reconstruction may prevent greater recall of phonologically similar than of phonologically dissimilar items. However, this does not necessarily imply that this will also produce a phonological similarity effect of lower recall of phonologically similar than phonologically dissimilar items. The reason why phonological similarity effects were absent previously may have been because previous phonological similarity effects in recall were actually similarity effects on the order of recall (Burgess & Hitch, 1992; Ellis, 1980; Gruneberg & Melton, 1972; Healy, 1974, 1982; Lewandowsky & Murdock, 1989; Watkins, et al., 1974; Wickelgren, 1965).

Upon close inspection of the data reported by Baddeley, et al. (1975) and Baddeley, et al. (1984), it was noted that the

stimuli used by those researchers were re-presented several times over each block of trials. Typically 5 words from a pool of 8 were used for up to 8 trials and so it can be argued that recall of these words was no more than recall of the order of the words. Also, as ample time to learn the words was given, it was mainly the order that the words occurred in that varied.

Whilst I was conducting Experiment 5, Coltheart (1993) published an article in which an almost identical design to Experiment 5 was used to examine whether sampling method affected the phonological similarity effect. Coltheart conducted two experiments varying across stimulus sampling method. Coltheart's first experiment examined the phonological similarity effect without articulatory suppression and her second experiment examined the phonological similarity effect with articulatory suppression. Coltheart also compared recall in the correct serial position with recall in any serial position. Both experiments presented 5 words, at one per second.

Coltheart's experiments differed in 2 ways from the present Experiment 5. First, Coltheart (1993) examined the effect of articulatory suppression using a *between subjects* design; the present Experiment 5 used a *within subjects* design. The trade-off in this issue is that Coltheart had more trials available to estimate recall per condition. In contrast, the present Experiment 5 had only one within subjects error term and hence the ability to *directly* estimate the interaction between phonological similarity and articulatory suppression. Second, Coltheart (1993) presented words at one per second; Experiment 5 of the present research allowed participant-paced presentation. While at this stage it is unclear what effect pacing method has on performance, this issue is addressed in Experiment 6.

Coltheart (1993) found that the phonological similarity effect was present with both a fixed and a non-replacement

pool sampling method. However, Coltheart's results did show that there were more order errors (as calculated in the present research) for phonologically similar than for phonologically dissimilar words. Also, in her first experiment, the results showed that there was a tendency toward greater statistical confidence for a phonological similarity effect produced from a fixed-pool than from a non-replacement pool. Thus, although Coltheart concluded that the phonological similarity effect is not solely an effect of sampling method, her results can also arguably support the hypothesis that repeated presentation of words produces a more robust effect of phonological similarity than lists with new items in each trial. More important, the increased number of order errors for phonologically similar items in Coltheart's results replicates the results of Experiments 1 and 3 of the present research. Experiment 5 of the present research was an independent replication of Coltheart's research in which the alteration of order errors and the phonological similarity effect due to pool-type were considered more carefully.

There is also evidence as to whether the Phonological Store is involved with order information to be found in research in where differences in some other phonological processing continue to be present in conditions where the Phonological Store is theoretically inactive. Specifically, the Phonological Store is presumed to be inactive when a person is using articulatory suppression (Baddeley, 1986). Therefore, if the Phonological Store is the repository of phonological information, then with articulatory suppression that information ought to be unavailable to a person. Using this reasoning, Besner and Davelaar (1982) argued that if articulatory suppression was preventing encoding of visual stimuli into a phonological code, then pseudo-homophones (e.g. *BRANE*) should be no better recalled than non-words (e.g., *FRANE*) under articulatory suppression. Whether words

were pseudo-homophones or non-words formed the factor of lexicality.

In their first experiment, Besner and Davelaar (1982) investigated the effect of phonological similarity, articulatory suppression, and lexicality (pseudo-homophones or non-words) on serial recall. Their data exhibited phonological similarity effects which were abolished by articulatory suppression and a general lower recall when persons were using articulatory suppression. These results were consistent with previous research (Baddeley, 1986, 1992a, 1992b). However, pseudo-homophones should have been no better recalled than non-words because articulatory suppression should have prevented the person accessing the base-word phonology of the pseudo-homophone. What Besner and Davelaar did find was that more pseudo-homophones were recalled than non-words. This increased recall of pseudo-homophones over non-words was not affected by articulatory suppression. Thus, it was inferred that the phonological equivalence of the pseudo-homophones to the base-words was available for processing, even after articulatory suppression had supposedly prevented encoding of visual stimuli to a phonological code. Besner and Davelaar inferred from these results that:

There are at least two phonological codes driven by print...The first code can be used for lexical access, and is not prevented from operating by suppression. The second phonological code is prevented from operating by suppression, and functions as a durable storage medium for retaining serial order information; this aids verbatim recall and comprehension. (p.708)

The problem with proposing two phonological stores is perhaps obvious; when a result occurs that does not fit the existing model, propose a new store. Whether two stores are necessary or not remains to be determined. What is important from

Besner and Davelaar is that their results were inconsistent with Baddeley's (1986) description of the Phonological Store and that they resolve this inconsistency by claiming that the original Phonological Store is *primarily* involved in storing order information. In Experiment 5, the present author chose to first examine whether the original Phonological Store does indeed store phonological information, or whether, as suggested by Besner and Davelaar, that store is primarily concerned with storing order information.

To determine whether the phonological similarity effect is in fact an effect of confusing the order of stimuli repeatedly presented, Experiment 5 directly compared recall for words chosen from a pool without replacement with recall for words chosen from a fixed pool with replacement.

From the above discussion, Experiment 5 aimed to first replicate a phonological similarity effect. The hypotheses necessary to adduce a phonological similarity effect were:

1.1 There would be greater recall of phonologically dissimilar than similar items;

1.2 There would be lower recall of all items when the person was using articulatory suppression; and

1.3 There would be no phonological similarity effect with concurrent articulatory suppression.

Experiment 5 also examined whether the phonological similarity effect was more robust when stimuli were sampled from a fixed pool with replacement in contrast to sampling from an item pool without replacement. This is an important issue, as if the phonological similarity effect is more robust in the fixed pool than the non-replacement pool condition, it could be advanced that the phonological similarity effect is more an effect of item order than of content confusion.

In summary, Experiment 5 investigated the phonological similarity effect in a logical manner designed to provide information sufficient to replicate it in a complex-span task in subsequent experiments.

Articulation rate

Having considered the contribution of Experiment 5 to the overall aim of integrating two models of working memory, the contribution of a high level articulatory production schema was also considered. It has been argued that the Articulatory Rehearsal Process of working memory and the generation of an articulatory motor code are produced by the same higher level process (Baddeley, et al., 1975; Gathercole & Baddeley, 1993). Therefore, sampling articulatory motor output would logically provide a secondary indicator of the higher level articulatory schema, and by inference of the operating rate of the Articulatory Rehearsal Process. The measure of the speed of the articulatory process used in the present research was the oral articulation rate of the items presented to the persons in the experimental trials (Baddeley, et al., 1975). The present Experiment considered whether articulatory rate might not also index the operation of the Phonological Store, a question that has remained virtually unexplored. If articulatory rate indexes both the Articulatory Rehearsal Process and the Phonological Store, then the argument for the fractionation of the Phonological Loop into a separate Articulatory Rehearsal Process and Phonological Store (Baddeley, 1986, 1992a, 1992b; Gathercole & Baddeley, 1993) would be demonstrably weakened.

Schweikert, et al. (1990) specifically investigated whether articulation rates were different for phonologically similar and dissimilar words. When each list to be recalled was presented, all consonants were on screen simultaneously. The person was to read all of the consonants on screen orally, and then to recall as many as they were able. The mean

duration to read each list was taken as the mean time to read each list. Schweikert, et al.'s data found that the mean time to read a list from 10 phonologically similar consonants (b, c, d, g, j, k, p, t, v, z) was 1.87 s and to read a list from 10 phonologically dissimilar consonants (b, d, f, h, l, k, m, q, r, z) was 2.42 s. When the mean reading durations were divided into the number of words recalled, the 'pronunciation rate' for similar words was 3.01 s and for dissimilar words was 2.92 s. Schweikert, et al. concluded that there were no differences in articulation rate between phonologically similar and dissimilar words. However, the time to read each list in this design was confounded with rehearsal of the order that the consonant occurred in. Thus, the measure of articulation rate used by Schweickert, et al. was not independent of the measure of recall. To obtain a purer measure of articulation rate against which to regress recall, it was considered methodologically more accurate to obtain an *independent* estimate of articulation rate rather than the arguably co-varying estimate used by Schweickert, et al.

A second potential confound in the Schweikert, et al. (1990) study was that the definition of what is phonologically similar or phonologically dissimilar appears to be different than in most research investigating the phonological similarity effect. There appears to be some degree of phonological similarity between the phonologically dissimilar letters. It seems to the present author that *b*, *d*, and *z* all contain the -ee phoneme. Also, *f*, *l*, and *m* all appear to share the same initial phoneme. Finally, *q* and *k* share the -k phoneme. Conversely, the 'phonologically similar' consonants do not all have one common phoneme as *j* and *k* do not share the -ee phoneme with the rest of the consonants.

In contrast, while preparing the program to present the stimuli for Experiments 5 and 6, it seemed to myself that phonologically similar words took longer to read than dissimilar words, what Caplan and Waters (1994) call

"'tongue-twister' effects" (p. 1057). After some informal trials with other persons, it was decided to formally investigate the relationship between recall and articulation rate for the phonological similarity effect in the same manner that the word-length effect was originally investigated (Baddeley, et al., 1975). Thus, at the end of Experiment 5, a complete list of all the words in each condition was presented to be read by each participant. So, with lists of phonologically similar repeating words, all of the words that were presented were in the list, in the order presented, and with repetitions present. (For example, if the stimuli used the word *chart* 3 times, *chart* was presented 3 times in the articulation rate list.) This design also allowed a comparison of whether articulation rates differed across sampling method. The hypothesis for articulation rate in Experiment 5 was that:

1.4 The articulation rate of phonologically similar and phonologically dissimilar words would be equivalent and *non-predictive* of the phonological similarity effect.

In summary, the importance of articulation rate to the phonological similarity effect is this: *If the phonological similarity effect is due to different rates of reading for similar than dissimilar words just as the word-length effect is due to different rates of reading depending on word length, then the Phonological Store cannot be as distinct from the Articulatory Rehearsal Process as first suggested.*

Order, processing time, and the phonological similarity effect

The second purpose of Experiment 5 was to further refine the measures of order and processing time, in addition to content. In Experiment 5, measures of order and processing time were investigated in relation to a phonological similarity effect on recall.

In considering the combined data from Experiments 1 to 4, a methodological weakness in how recall in the correct position was assessed also became apparent. In using the two measures of content in the first 4 studies in this research program, no check was made on whether the measure of recall in the correct serial order was a guess or whether the person was even relatively certain that the word was in that position. For example, if a person correctly wrote the words in position 1, 2, 3, and 5, two possibilities existed. The person may have been sure that each of the four words were indeed in the positions indicated. Alternately, the person may have guessed the positions of one or more of the words. If the person was sure of 3 out of 4 positions, then they would have a $1/3$ chance of correctly guessing position 4. This scenario is not an unreasonable one as most people recalled around 3 to 4 words in the correct serial position. Thus, to the extent that the score of words recalled in the correct position has this guessing component, the score is an over-estimate of recall in the correct serial position. Concomitantly, the effect of order, which is the recall in any order minus recall in the correct serial position, is probably an under-estimate of the true effect of order. Clearly, the solution of asking for an estimate of the number of words the person is certain are in the correct position is a useful thing to do, and will be done in succeeding research. So, in Experiment 5 a post-estimate of recall in the correct serial position was included so that it could be compared to performance measures to test whether persons were guessing many serial positions of items. Persons could be inferred to be guessing serial position of items if they reported significantly lower post-estimates than their actual recall in the correct serial position.

Finally, Experiment 5 further pursued the developing hypothesis that processing time at input is separate from not only the word-length effect, but from the phonological similarity effect also. That is, based on Experiments 1 to 4,

processing time appears to be unrelated to subsequent performance. This is also despite the rather variable times taken by persons to view the stimuli on some trials.

In Experiments 1 to 4, order information produced data that appeared to differentiate between effects resulting from the phonological similarity of items and effects resulting from item length. The greater effect of phonological similarity upon order errors for phonologically similar than for phonologically dissimilar items was hoped to be replicated in Experiment 5. So, in Experiment 5, it was expected that:

2.1 phonologically similar items would produce more order errors than phonologically dissimilar items.

Also, in Experiments 1, 2, and 4, data showed that articulatory suppression had no effect on order information. In Experiment 3, articulatory suppression *increased* the loss of order information over when no articulatory suppression was used. Generally, it appears that articulatory suppression has little or no effect on order information. Experiment 5 sought to further elucidate whether articulatory suppression affects order information or not. If it does not, then there is support that the loss of order information relates to loss of information from a Phonological Store and not from a failure of the Articulatory Rehearsal Process. The hypothesis of the present research was that:

2.2 If order errors are not produced by the Articulatory Rehearsal Process, then there should be no effect of articulatory suppression on order errors.

Data from Experiments 1 to 4 have been interpreted as showing that processing time at input may have no effect on either the Phonological Store or the Articulatory Rehearsal Process (both maintenance and output processes). This conclusion was clearly testable in Experiment 5. If processing time at input

has no effect on maintenance or output processes then it was expected that:

2.3 viewing time would be unrelated to any phonological similarity effects on recall.

Meta-memory and the phonological similarity effect

The third purpose of Experiment 5 was to investigate participants' ability to estimate their level of recall before they perform the working memory task. But more important, Experiment 5 again provided the opportunity for a direct test of persons' ability to monitor the phonological similarity effect when it was also present in a performance measure.

One improvement in how on-line meta-memory was measured in Experiment 5 was that participants were given information prior to each trial as to the nature of the task and after they had learned in the practice trials what the descriptions of the conditions meant (e.g., "This trial uses similar words with counting"). In Experiments 1 to 4, persons knew that conditions were presented in three-trial groups. However, they were unaware of what the next group would be. So, any pre-estimates involved a guess as to the nature of the first trial of each group. Despite this flaw, participants still demonstrated an impressive ability to monitor their working memory. Thus, in Experiment 5, each pre-estimate was a more informed pre-estimate than in Experiments 1 to 4. From this information and from experience on the practice trials, each person made a numeric estimate of their expected level of recall.

Data from Experiments 1 to 4 supported that content measures of on-line meta-memory mirror content measures of memory performance in pattern and are related to each other in a regular fashion. However, at this point the present research not only aimed to replicate previous findings in Experiments

1 to 4, but to directly examine if persons could predict the pattern of the phonological similarity effect, assuming there is one, in Experiment 5. In addition to predicting the pattern of the phonological similarity effect, Experiment 5 also aimed to establish a statistical relationship between on-line meta-memory pre-estimates and working memory recall. Because establishing a link between a phonological similarity effect in performance and pre-estimates is without precedent, no specific hypothesis was advanced.

Summary

Experiment 5 was principally concerned with investigating the parameters for replication of the phonological similarity effect. Specifically, it was hypothesized that the time available at recall and factors affecting the amount of order information to be recalled were salient parameters in producing a phonological similarity effect. In addition to a measure of content, Experiment 5 investigated the relevance of measures of order and processing time in conditions designed to replicate the phonological similarity effect. Finally, Experiment 5 investigated whether persons were able to predict their working memory performance in general and, more important, whether persons could predict a phonological similarity effect.

Method

The method for Experiment 5 followed the general method detailed previously. Only the specific differences from the general structure are presented below.

Participants

Participants were student volunteers recruited from Student Job Search and were paid \$5.00 for their participation. There were 34 people in Experiment 5; 18 in the group who received stimuli from the fixed pool sampled from with replacement

(Mean age = 22.83, SD = 3.93) and 16 in the group who received stimuli sampled from without replacement (Mean age = 22.37, SD = 3.54). There was no age difference between groups, $t(32) = .36$, $p = .724$, two-tailed. Due to a computer error, one participant's data was lost for their pre-estimates, reducing the sample size for that dependent variable from $N = 34$ to $N = 33$.

Design

The design was a three-factor mixed design. The factors of word-type (phonologically similar, phonologically dissimilar) and articulatory suppression (control, articulatory suppression present) were varied within subjects. The factor of pool-type (fixed, non-replacement) was varied between subjects. The dependent variables were as detailed in the General Method section (p. 99).

Procedure

Prior to arrival, each participant was randomly placed in either the Fixed-pool or Non-replacement pool condition (p. 91). Otherwise the procedure was as detailed in the General Method section (p. 95).

Results

The following results present data showing that recall is lower when stimuli are phonologically similar than when phonologically dissimilar. That is, a phonological similarity effect is present in Experiment 5. The data also supported the premise that the phonological similarity effect is produced because of loss of order information. In addition, the data demonstrated the independence of processing time at input from the phonological similarity effect. And finally, data showed that persons were unable to monitor working memory sufficiently to predict the phonological similarity effect.

Content measures: Recall in the correct serial position and recall in any serial position

Recall in the correct serial position

Overall, whether words were from a fixed or non-replacement pool had no effect on recall in the correct serial position. That is, there were no main or interaction effects with pool-type.

Phonologically similar words were less well recalled than phonologically dissimilar words, $F(1,32) = 43.83$, $p < .001$, $MSE = .26$, in the control condition, as predicted by Hypothesis 1.1. The effect of articulatory suppression was to lower recall irrespective of word-type or pool-type, $F(1,32) = 98.33$, $p < .001$, $MSE = .41$, confirming Hypothesis 1.2. When persons were using articulatory suppression, recall in the correct serial position showed no phonological similarity effects. The removal of the phonological similarity effect with articulatory suppression produced a word-type \times articulatory suppression interaction, $F(1,32) = 98.33$, $p < .001$, $MSE = .41$ (Table 24; Figure 9), as predicted by Hypothesis 1.3.

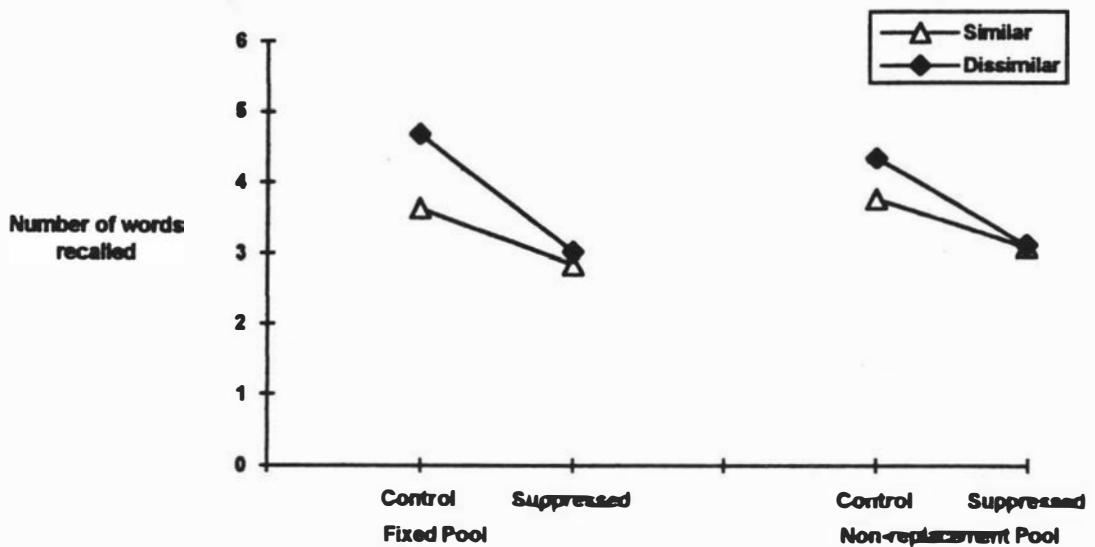


Figure 9. Recall in the correct serial position by phonological similarity, articulatory suppression, and pool-type conditions in Experiment 5.

In summary, Experiment 5 replicated the phonological similarity effect in a simple-span task. Second, using the measure of recall in the correct serial position, there was no statistical basis to assert that sampling from a fixed size pool with replacement, as opposed to from an unlimited pool without replacement, differentially affected the presence of a phonological similarity effect.

Finally, when post-estimates of recall in the correct serial position were compared with performance data, there were significant correlations between each measure. Also, there were no significant differences in magnitude between performance and post-estimate measures. Taken together, these results suggest that persons were not guessing item positions and that the performance measures probably represented a valid measure of how many item positions were recalled (Table 24).

Recall in any serial position

Recall in any serial position showed a phonological similarity effect (in the control condition) as predicted by Hypothesis 1.1, $F(1,32) = 8.68$, $p = .006$, $MSE = .09$. That is, fewer phonologically similar words were recalled than phonologically dissimilar words. It was found (Hypothesis 1.2) that articulatory suppression lowered recall overall, $F(1,32) = 86.34$, $p < .001$, $MSE = .33$, and that with articulatory suppression there was no longer a phonological similarity effect (Hypothesis 1.3). The absence of the phonological similarity effect with articulatory suppression where it was present in the control condition produced a word-type x articulatory suppression interaction, $F(1,32) = 14.07$, $p = .001$, $MSE = .12$ (Figure 10).

Unlike recall in the correct serial position, recall in any serial position had a phonological similarity x pool-type interaction, $F(1,32) = 29.50$, $p < .001$, $MSE = .18$. This interaction resulted from the greater recall of phonologically dissimilar than of phonologically similar

items in the fixed pool condition, $F(1,17) = 40.16$, $p < .001$, $MSE = .06$, but not in the non-replacement pool condition. These data (Table 24; Figure 10) suggest that sampling from a fixed length pool with replacement produces a more robust phonological similarity effect than sampling from an unlimited pool without replacement.

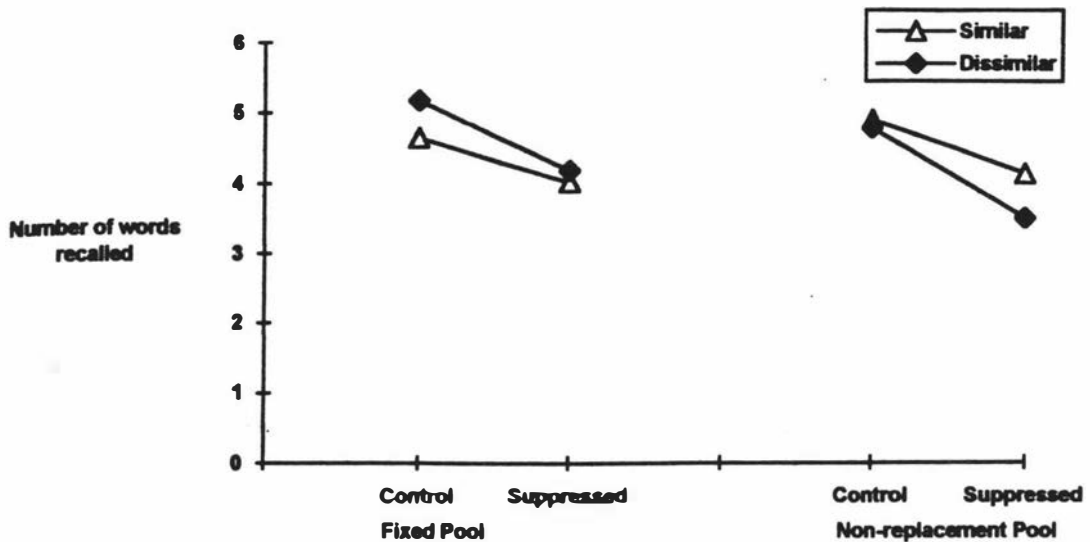


Figure 10. Recall in any serial position by phonological similarity, articulatory suppression, and pool-type conditions in Experiment 5.

Summary

When stimuli were drawn from a fixed pool with replacement, both measures of content showed a phonological similarity effect without but not with articulatory suppression. In contrast, for stimuli drawn from a non-replacement pool, only recall in the correct serial position displayed a phonological similarity effect. Thus, phonological similarity effects were more consistently produced from stimuli drawn from a fixed than from a non-replacement pool.

It has also been advanced that the phonological similarity effect is an effect of the loss of order information. If this is so, then one would expect to find a larger phonological similarity effect for recall in the correct serial position, where order information is necessary, than for recall in any

serial position. An illustration will help present this important point. If a person recalled 4 out of 6 phonologically dissimilar and 3 out of 6 phonologically similar words in the correct serial position, then there is a $3/4 = 0.75$ ratio of phonologically similar to phonologically dissimilar words. If the person's recall in any serial position increases to 5 out of 6 phonologically dissimilar words, then if the phonological similarity effect is due to loss of content alone, recall of phonologically similar words ought to be about $0.75 \times 5 = 3.75$ words. In this example, an interaction for absolute values of recall would occur such that the phonological similarity effect would increase with recall in any serial position over recall in the correct serial position.

Table 24. Mean (SD) recall of phonologically similar or phonologically dissimilar words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions and across pool-type in Experiment 5 ($N = 34$).

	Measures of Content		
	Recall in the correct serial position	Recall in any serial position	Post-estimates of recall in the correct serial position
	Fixed pool-type		
Control			
Similar	3.64 (1.05)	4.64 (0.51)	3.74 (0.80)
Dissimilar	4.70 (0.82)	5.18 (0.64)	4.46 (0.98)
Articulatory suppression			
Similar	2.84 (1.15)	4.00 (0.97)	2.90 (0.96)
Dissimilar	3.02 (1.04)	4.17 (0.81)	3.08 (0.71)
Non-replacement pool-type			
Control			
Similar	3.76 (1.05)	4.87 (0.62)	4.08 (1.00)
Dissimilar	4.34 (1.01)	4.75 (0.74)	4.32 (1.10)
Articulatory suppression			
Similar	3.08 (0.97)	4.10 (0.73)	3.28 (1.18)
Dissimilar	3.12 (0.98)	3.47 (0.85)	3.14 (1.18)

In Experiment 5, the data show that the size of the phonological similarity effect decreased from recall in the correct to recall in any serial position, $F(1,32) = 30.97$, p

$< .001$, $MSE = .10$ (Figure 11). This interaction suggests that the phonological similarity effect is more robust as more order information is required in recall. That is, the phonological similarity effect appears to be produced, at least in part, by the differential loss of order information as a function of phonological similarity.

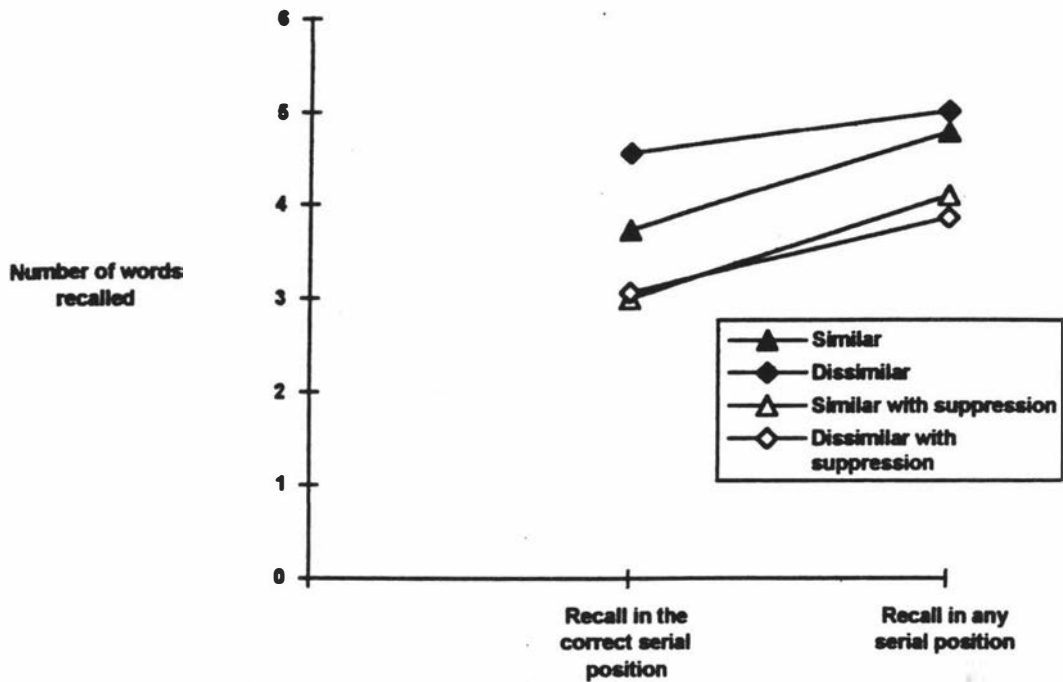


Figure 11. A comparison of the two measures of content in Experiment 5 with each other.

Articulation rate and its effect on recall of phonologically similar and dissimilar words

Experiment 5 was the first experiment in the present dissertation to include a measure of articulatory activity. The measure of articulatory activity used in Experiment 5 was articulation rate; articulation rate is the mean number of words read aloud per second from each word-type condition (Table 25).

There was no effect of pool-type on articulation rate, but both pool-type \times word-type, $F(1,32) = 24.90$, $p < .001$, $MSE = .04$, and pool-type \times word-type \times articulatory suppression,

$F(1,32) = 18.98$, $p < .001$, $MSE = .01$, interactions were present. Explanation of the meaning of these interactions is best done through considering each pool-type separately.

Table 25. Articulation rate (SD) for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions and across pool-type in Experiment 5 (N = 34).

	Fixed pool		Non-replacement pool	
	Control	Suppressed	Control	Suppressed
Similar	1.84 (0.37)	1.68 (0.37)	2.03 (0.47)	2.01 (0.43)
Dissimilar	2.11 (0.41)	2.14 (0.44)	2.05 (0.48)	1.99 (0.40)

For the Non-replacement pool, articulation rates did not differ as a function of articulatory suppression or word-type, and nor were there any interactions. This important result confirms that a phonological similarity effect can occur without a difference in articulation rate. That is, a difference in articulation rate is not a necessary condition to a phonological similarity effect, whereas it is for a word-length effect (Baddeley, 1986, 1992a, 1992b).

For the Fixed pool, there was no difference between the articulatory suppression and control conditions. However, in contrast to the Non-replacement pool, articulation rate was slower for phonologically similar than for phonologically dissimilar words in the Fixed pool condition, $F(1,17) = 42.27$, $p < .001$, $MSE = .06$. Finally, there was an interaction between articulatory suppression and word-type articulation rates, $F(1,17) = 28.13$, $p < .001$, $MSE = .01$. This interaction was examined with post-hoc paired t-tests. In both the control list, $t(17) = -4.91$, $p < .001$, one-tailed, and the articulatory suppression list, $t(17) = -7.37$, $p < .001$, one-tailed, phonologically similar words took longer to articulate than phonologically dissimilar words. However, the phonologically similar words in the control list were also articulated slightly faster than in the articulatory suppression list, $t(17) = 4.21$, $p = .001$, two-tailed, thus

producing the articulatory suppression x word-type interaction. The consequences of this interaction for the experimental results are nil because if words that were recalled with articulatory suppression had a slower articulation rate than words without articulatory suppression, a greater difference in recall for articulatory suppression would have been predicted. Instead, there was lower recall for words recalled with articulatory suppression.

However, not only must there be both lower articulation rates and lower recall for phonologically similar than for phonologically dissimilar words without articulatory suppression, but there must also be some demonstrated relationship between these measures (Hypothesis 1.4). The question is now whether articulation rate was related to recall in the Fixed-pool condition. One problem with regressing recall against articulation rate was that recall measures had a larger variance than the articulation rate measures which could have reduce the linearity of the regression term (Keppel, 1982, p.557). It was decided to use regression equations with natural logarithm transformations of both storage and articulation rate measures. This transformation process improved the fit of the regression curves (Table 26).

Table 26. The proportion of variability accounted for in the Fixed-pool condition using the natural logarithm of each of the content measures regressed against the natural logarithm of articulation rate with *F*-values, significance, and mean square errors of the residual terms evaluating the fit of the regression coefficient.

	<i>R</i> ²	<i>F</i> -value	significance	<i>MS</i> _{residual}
Recall in the correct serial position				
Similar	.17	3.22	.092	.09
Dissimilar	.21	4.21	.057	.03
Recall in any serial position				
Similar	.32	7.45	.015*	.01
Dissimilar	.27	5.92	.027*	.01

* *p* < .05

For recall in the correct serial position, both phonologically similar and dissimilar words showed a trend toward significant prediction of recall by knowing the articulation rate of those words. For recall in any serial position, the regression coefficients for both phonologically similar and phonologically dissimilar words accounted for 27%-32% of the variability in recall given only information about articulation rate.

In summary, depending on what measure of recall was employed, from 27%-32% of the actual recall could be predicted from information about articulation rate alone when words were from a small pool with replacement. Thus, the assumption that the phonological similarity is unrelated to articulation rate was in part disconfirmed in Experiment 5. That is, because persons got more 'tongue-tied' (Caplan & Waters, 1994) with phonologically similar than phonologically dissimilar words, their slower rate of articulation may have contributed toward the phonological similarity effect. This relationship between articulation rate and recall in the fixed pool condition might provide an explanation of why, when stimuli were repeatedly presented, that the phonological similarity effect became larger. That is, order confusions *plus* slower articulation for phonologically similar items created a larger phonological similarity effect in the fixed pool condition than in the non-replacement condition.

In conclusion, Hypothesis 1.4, which stated that the articulation rate of phonologically similar and phonologically dissimilar words would be equivalent and non-predictive of the phonological similarity effect, was only partially supported. The data in Experiment 5 showed that a phonological similarity effect might be increased in size by slower articulation rates, but that a phonological similarity effect can indeed exist without a difference in articulation rate as Baddeley (1986, 1992a, 1992b) originally described.

The effect of Order

order errors were again calculated as the number of words recalled in any serial position minus the number of words recalled in their correct serial position (Table 27). Overall, pool-type had no effect on order errors. However, both phonological similarity differences, $F(1,32) = 10.36$, $p = .003$, $MSE = .19$, and articulatory suppression differences, $F(1,32) = 8.75$, $p = .006$, $MSE = .17$, interacted with pool-type. Consequently, order errors were analyzed separately for each pool-type.

Table 27. Mean (SD) order errors for phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions and across pool-type in Experiment 5 ($N = 34$).

	Pool-type	
	Fixed ($N = 18$)	Non-replacement ($N = 16$)
Control		
Similar	1.00 (0.71)	1.01 (0.60)
Dissimilar	0.48 (0.28)	0.41 (0.39)
Articulatory suppression		
Similar	1.16 (0.63)	1.02 (0.45)
Dissimilar	1.15 (0.66)	0.35 (0.25)

For the Fixed pool condition, there were more order errors for phonologically similar than for phonologically dissimilar words, $F(1,17) = 7.78$, $p = .013$, $MSE = .16$, thus supporting Hypothesis 2.1. There were more order errors with articulatory suppression than in the control condition, $F(1,17) = 16.74$, $p = .003$, $MSE = .18$, which disconfirms Hypothesis 2.2. There was also an interaction between word-type and articulatory suppression, $F(1,17) = 5.34$, $p = .034$, $MSE = .22$. Post-hoc one-tailed paired t -tests showed the nature of the interaction was such that in the control condition there were more order errors for phonologically similar than phonologically dissimilar words, $t(17) = 3.69$, $p = .001$. However, in the articulatory suppression condition

there was no difference in order errors between word-types, $t(17) = .06$, $p = .476$.

For the Non-replacement pool condition, an increase in order errors was found for phonologically similar over phonologically dissimilar words, $F(1,15) = 41.63$, $p < .001$, $MSE = .18$, supporting Hypothesis 2.1. There was no change in the level of order errors with articulatory suppression supporting Hypothesis 2.2.

To summarize, order errors were greater for phonologically similar than for phonologically dissimilar words without articulatory suppression, in both the Fixed and Non-replacement pools, as predicted. That is, results from Experiment 5 support a conclusion that the phonological similarity effect is produced by order errors. To further confirm this conclusion the present author decided to scrutinize the nature of the order errors by examining item transpositions (e.g., Healy, 1974, 1982). Whereas order errors examined the difference between recall in the correct and in any serial position, item transpositions show how large and in what serial positions those order errors occur. Item transpositions represent the frequency with which each item was transposed to another serial position and the distance from its original serial position (Figure 12).

As can be seen in Figure 12, item transpositions were typically close to the correct serial position and more item transpositions occurred in the middle than the outer serial positions for both phonologically similar and for phonologically dissimilar items. Thus, the qualitative nature of item transpositions was equivalent for both phonologically similar and phonologically dissimilar items. However, what is important in Experiment 5 is that there were more item transpositions for phonologically similar than for phonologically dissimilar items, $\chi^2(1, N=34) = 36.63$, $p = .001$. This implies that the phonological similarity effect is produced by a quantitatively greater number of item

transpositions for phonologically similar than phonologically dissimilar items, but that the qualitative nature of transpositions remains the same. That is, as was shown with analysis of order errors, the *phonological similarity effect* appear to be an effect produced by loss of item position information.

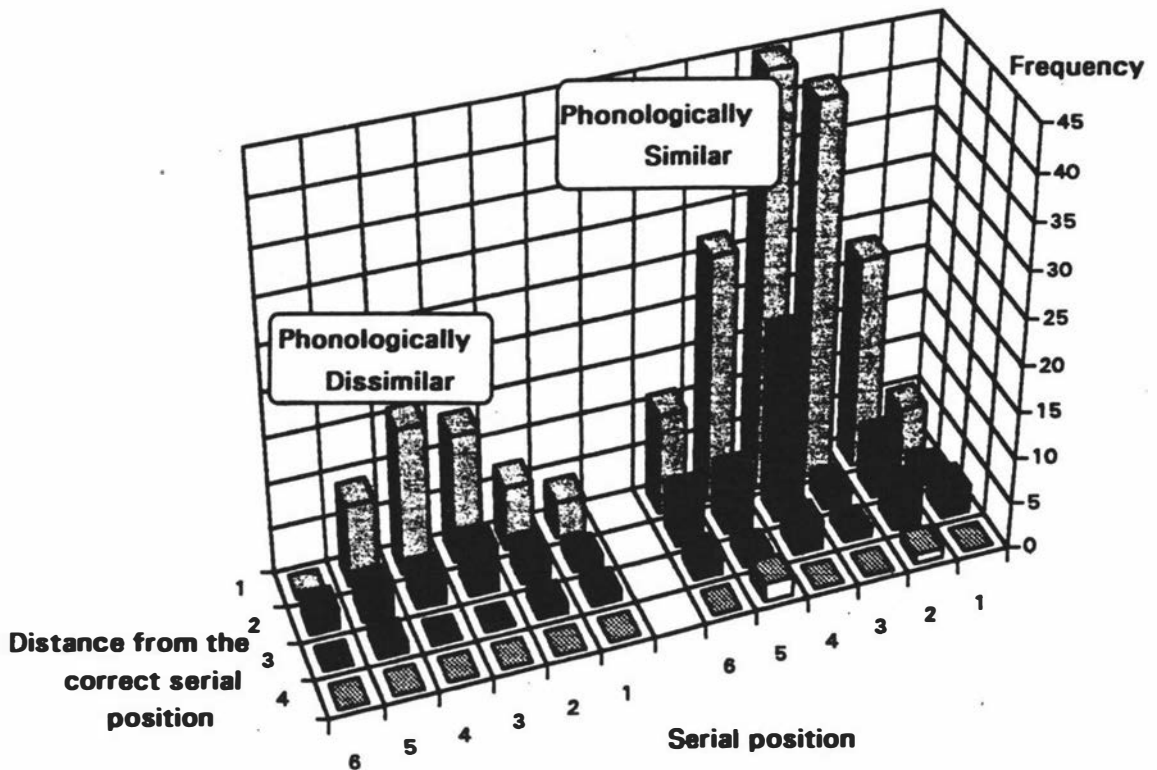


Figure 12. Item transpositions for phonologically similar and phonologically dissimilar words in Experiment 5.

Processing time: Viewing time

Overall there were no effects of pool-type, word-type, or articulatory suppression for the viewing time measure. There was a word-type x articulatory suppression interaction, $F(1,32) = 12.11$, $p = .001$, $MSE = 35,856$. This interaction occurred because phonologically similar words were viewed for longer than dissimilar words, $F(1,32) = 6.88$, $p = .013$, $MSE = 77,106$, in the control condition, contrary to Hypothesis 2.3. However, there was no significant difference in viewing times

in the articulatory suppression condition between phonologically similar and dissimilar words (Table 28).

However, not only must viewing time be different across phonological similarity, but there must be some relationship between it and recall in order to infer some causality. A regression analysis was performed to examine if any of the content measures had a significant amount of variability predicted by the viewing time of the stimuli. The regression of content onto viewing time showed that no measures of content were significantly related to viewing time. Overall, the data show that viewing time as an input process is unconnected with the output and maintenance processes as measured by recall. That is, the data are consistent with previous data in the present dissertation that have also shown the dissociability of the input process from the phonological similarity effect.

Table 28 Mean (SD) viewing times (ms) for both phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions and across pool-type in Experiment 5 (N = 34).

	Pool-type	
	Fixed	Non-replacement
Control		
Similar	1751 (863)	1930 (1336)
Dissimilar	1523 (568)	1803 (1074)
Articulatory suppression		
Similar	1630 (569)	1828 (1247)
Dissimilar	1662 (575)	1894 (1139)

In summary, while viewing times were longer for phonologically similar than dissimilar words without articulatory suppression, these slightly longer viewing times had negligible influence on content measures.

On-line meta-memory: Pre-estimates of recall in any serial position

There was no effect of pool-type on pre-estimates, but there were interactions between pool-type and word-type, $F(1,31) =$

7.38, $p = .011$, $MSE = .14$, and between pool-type, word-type, and articulatory suppression, $F(1,31) = 5.46$, $p = .026$, $MSE = .09$ (Table 29). Explanation of the meaning of these interactions is best done through considering each pool-type separately.

For the Fixed pool, there was an overall decline in pre-estimates due to articulatory suppression, $F(1,17) = 122.45$, $p < .001$, $MSE = .11$. However, there was no difference in pre-estimates between phonologically similar and phonologically dissimilar words. That is, persons were unable to predict the pattern of recall from information about phonological similarity, but were able to predict the effect of articulatory suppression on their performance. Yet, for the control condition, there was a moderate relationship between pre-estimates and performance for both phonologically similar, $R^2 = .37$, $F(1,16) = 9.56$, $MS_{residual} = 0.17$, and phonologically dissimilar conditions, $R^2 = .46$, $F(1,16) = 13.51$, $MS_{residual} = 0.24$ at $p < .05$. Also, there was a moderate relationship between pre-estimates and performance measures for phonologically similar and for phonologically dissimilar words in the articulatory suppression condition, $R^2 = .40$, $F(1,16) = 10.58$, $MS_{residual} = 0.24$, and , $R^2 = .32$, $F(1,16) = 7.42$, $MS_{residual} = 0.48$, respectively, at $p < .05$.

Table 29. Mean pre-estimates (SD) for both phonologically similar and phonologically dissimilar words for both the control and articulatory suppression conditions and across pool-type in Experiment 5 (N = 33).

	Pool-type	
	Fixed	Non-replacement
Control		
Similar	3.77 (0.76)	4.37 (0.96)
Dissimilar	3.91 (0.96)	3.91 (1.06)
Articulatory suppression		
Similar	2.96 (0.81)	3.29 (0.97)
Dissimilar	2.95 (0.70)	3.17 (1.03)

For the Non-replacement pool, again there was an overall decline in pre-estimates due to articulatory suppression, $F(1,14) = 42.05$, $p = .001$, $MSE = .30$, as predicted by Hypothesis 3.1. There was also a higher mean pre-estimate for phonologically similar than for phonologically dissimilar words, $F(1,14) = 6.16$, $p = .026$, $MSE = .20$, in the non-replacement pool-type which produced the pool-type \times word-type interaction. There was also a word-type \times articulatory suppression interaction for the non-replacement pool, $F(1,14) = 7.64$, $p < .001$, $MSE = .05$, which did not occur in the fixed pool, thus explaining the pool-type \times word-type \times articulatory suppression interaction. The word-type \times articulatory suppression interaction was examined using post-hoc paired t-tests. Pre-estimates were higher for phonologically similar than dissimilar words, $t(14) = 3.18$, $p = .007$, two-tailed, in the control condition, but were not significantly different, $t(14) = 1.04$, $p = .317$, two-tailed, with articulatory suppression (Table 29). Again, persons were able to predict the effect of articulatory suppression but not of phonological similarity.

In considering the relationship of pre-estimates to performance, again in the control condition there was a moderate relationship for phonologically similar, $R^2 = .38$, $F(1,13) = 8.09$, $MS_{residual} = 0.25$, and phonologically dissimilar words, $R^2 = .31$, $F(1,13) = 5.73$, $MS_{residual} = 0.03$, respectively, at $p < .05$. There was also a statistically significant relationship at $p < .05$ between pre-estimates and performance measures for phonologically similar and for phonologically dissimilar words in the articulatory suppression condition, $R^2 = .41$, $F(1,13) = 9.19$, $MS_{residual} = 0.01$, and $R^2 = .61$, $F(1,13) = 20.55$, $MS_{residual} = 0.001$, respectively.

In summary, persons were not able to predict the pattern of their recall given only phonological similarity information for either pool. This was despite there being a significant

relationship in each instance between pre-estimates and performance. In contrast, persons were able to predict the detrimental effect of articulatory suppression on recall in any serial position. So, in the first direct test of persons' ability to predict the effect of phonological similarity on their working memory performance, the data showed that they were simply not sensitive enough to do so.

Discussion

The present study examined three areas: the parameters necessary to replicate the phonological similarity effect; the use of measures of storage, processing time, and order information; and the effectiveness of on-line meta-memory.

Experimental parameters necessary for a phonological similarity effect

Experiment 5 demonstrated a phonological similarity effect, whereas Experiment 1 (a complex-span task) and Experiment 3 (a simple-span task) did not. However, Experiment 5 differed from Experiments 1 and 3 in two critical ways: Experiment 5 limited the recall time for words to prevent reconstruction and Experiment 5 examined the effect of sampling from a fixed pool of words. Thus the conclusion to be taken from Experiment 5 is two-fold regarding the experimental parameters necessary to replicate the phonological similarity effect. The phonological similarity effect requires the time for recall to be limited to prevent reconstruction of items from a common root. Limiting recall time prevented greater recall of phonologically similar than phonologically dissimilar items.

Second, Experiment 5 found that items sampled from a small pool with replacement produced a more robust phonological similarity effect than items sampled from a pool without replacement. The implication to be taken from this is that the phonological similarity effect is best replicated by

repeatedly presenting items drawn from a small pool rather than presenting novel items for each trial.

It must be acknowledged that the effect of pool-type was far from massive. When the conclusion from Experiment 5 that sampling method has an effect on recall is contrasted with Coltheart's (1993) conclusion that sampling does not have an effect, it is clear that this issue is in need of further investigation. However, it is to be borne in mind that the main purpose of Experiment 5 was to establish the conditions likely to maximize the phonological similarity effect. This purpose was achieved. Also, at the heart of the issue of sampling method is whether the phonological similarity effect is produced by successive order errors or by a loss of content. The present set of studies do not pursue the issue of sampling method, an issue of methodology. Instead, the issue of whether the phonological similarity effect is produced by order errors is examined by other, perhaps more germane, methods.

Finally, it has been assumed in the past that the phonological similarity effect was unrelated to the articulation rate of phonologically similar and dissimilar words (Schweickert, et al., 1990). Overall, the data in Experiment 5 were consistent with this assumption that the phonological similarity effect is not produced by articulation rate differences. However, data from Experiment 5 did demonstrate some relationship between articulation rate and phonological similarity effects in the Fixed-pool condition. However, the relationship does not appear strong enough to produce a phonological similarity effect, although it is possible that the increased articulation time for phonologically similar items in the fixed pool may have increased the magnitude of the phonological similarity effect.

Measures of order and processing time

order errors also show a clear difference in magnitude depending on the phonological similarity of the words used. There were more order errors for phonologically similar than for phonologically dissimilar words. Also, the detailed analysis of transposition errors confirmed that item positions of phonologically similar words were confused more frequently than of phonologically dissimilar words. However the nature of the transposition errors remained the same (see Healy, 1974, 1982 for similar analyses and conclusions). These results suggest that the phonological similarity effect is produced because more order information is lost with phonologically similar than with phonologically dissimilar words, even though the content is maintained (Gruneberg & Melton, 1972; Watkins, et al., 1974; Wickelgren, 1965).

From Experiment 5, there is now mounting evidence that the phonological similarity effect results from degradation of an order preservation process. This order-preservation process may be similar to the process of addressing information used in software. In a simple model of computer architecture, each item must have a reference (pointer) to the previous item and to the subsequent item. If one item is forgotten, then the pointer moves to the next item. This simple process allows items to be recalled in order but also allows some items to be forgotten. However, if each item has a constant number of pointers, then the effect of order on recall should be a constant proportion of actual recall, as has been empirically shown in previous studies. While the assertion that the phonological similarity effect represents degradation of an order maintaining process is novel in terms of the explicated theory of the qualitative/structural model of working memory (Baddeley, 1986, 1992a, 1992b), there are other descriptions of the phonological similarity effect (e.g., Conrad, 1964, 1965; Healy, 1974, 1982) and of working memory (e.g., Lewandowsky & Murdock, 1989; Schneider, 1993; Schneider &

Detweiler, 1987) that do predict such a process. In fact, Healy (1982) specifically describes this loss of order information as one of loss of *temporal* item information in comparison to loss of *spatial* information. The role of order information in working memory will be discussed in more detail in the General Discussion section (below).

Finally, Experiment 5 shows that more time was typically spent viewing phonologically similar than phonologically dissimilar words. However, this increased processing time at input appears unrelated to subsequent recall of items. It was as if, provided the upper and lower limits of the duration of working memory were not exceeded, the time taken to view stimuli had no influence on recall of those stimuli. These results are consistent with the argument of this dissertation that input processes are separate from maintenance/output processes.

On-line Meta-memory

Experiment 5 provided further evidence that people are able to accurately predict the effect of articulatory suppression on their performance. In contrast, persons were unable to predict the impact of phonological similarity on subsequent recall, with the inference that persons are thus unable to monitor working memory changes as a function of phonological similarity. The importance of this result is that, if persons cannot monitor the impact of phonological similarity, then, to the extent that a phonological similarity effect indicates operation of the Phonological Store, then one might infer that persons are perhaps insensitive to the Phonological Store. In terms of Postulate 3, the present results tend to disconfirm that persons can directly monitor the Phonological Store of working memory.

EXPERIMENT 6: A COMPARISON OF THE EFFECT OF EXPERIMENTER-PACED VERSUS PARTICIPANT-PACED PRESENTATION OF STIMULI ON WORD-LENGTH EFFECTS.

Experiment 6 follows from Experiment 5 in that it examined the parameters necessary for replication of the word-length effect (Figure 3, p. 87). Second, Experiment 6 continued to examine measures of order and processing time. Third, provided the word-length effect could be replicated, Experiment 6 aimed to examine persons' meta-memory for word-length effects.

Parameters for replicating the word-length effect

There were two probable reasons advanced in discussion of Experiments 2 and 4 as to why the word-length effect has been absent so far. These reasons are tested in Experiment 6. First, the difference in syllable length between the two word length conditions may have been insufficient in Experiments 2 and 4 -- an issue of experimental sensitivity. In Experiment 6, word lengths are 1 and 3-syllables compared to 1 and 2-syllables previously. Second, it may be that the time persons spend viewing the stimuli may affect the word-length effect. Although the present dissertation and other research (e.g., Avons, et al., 1994; Caplan, et al., 1992; Caplan & Waters, 1994; Coltheart, et al., 1990; Cowan, 1992; Cowan, et al., 1992; Henry, 1991; Monsell, 1987; c.f., Baddeley, & Andrade, 1994) suggests that rehearsal (input) and word-length (output) processes are separate, Experiment 6 explicitly examines the hypothesis that word-length effects are more robust when viewing time is experimenter-paced rather than participant-paced.

Experimental sensitivity and item discriminability

Experiments 2 and 4 both failed to find a word-length effect, but did demonstrate an effect of articulatory suppression. In Experiments 2 and 4, for complex-span and simple-span measures of working memory, respectively, it was predicted that word-length would influence recall. The failure of

Experiments 2 and 4 to find a word-length effect was surprising. To help understand why the word-length effect was not been replicated, a meta-analysis of Experiments 2 and 4 was conducted. (The data from Experiments 1 and 3 were combined using SPSS.PC's JOIN command.) A new variable GROUP was formed to distinguish between the simple and complex variants used in Experiments 2 and 4 respectively. Joining the data was justifiable because the sole difference between Experiments 2 and 4 was the change from sentences plus words to words alone.

When data from Experiments 2 and 4 were examined together more 1-syllable words (Mean = 4.45, SD = 1.35) were recalled than 2-syllable words (Mean = 4.20, SD = 1.21), thus showing a meagre word-length effect, $t(39) = 1.80$, $p = .04$ one-tailed in the control condition for recall in the correct serial position. In contrast, post-hoc one-tailed t-tests showed no effect of word-length for recall in any serial position. It was appropriate in this instance to use a one-tailed t-test because the direction of the expected and actual effects were known and were theoretically based. Given that in Experiments 2 and 4 that in all instances the 2-syllable words tended to be less well recalled than the shorter 1-syllable words and that the meta-analysis shows some meagre word-length effects, it was considered worth examining the statistical sensitivity of the design. An analysis of power is one way to achieve this (Keppel, 1982).

Analysis of statistical power for the combined data for Experiments 2 and 4 confirmed that there was low experimental sensitivity for the alternate hypotheses in those experiments; Experiments 2 and 4 showed less than 40% confidence in accepting the alternate hypotheses of better recall of one than two-syllable words and of a word length by articulatory suppression interaction at $p < .05$. In conclusion, Experiments 2 and 4 showed low experimental

sensitivity as a probable reason for the absence of a clear word-length effect.

Keppel (1982, p.528) suggests three reasons for low experimental sensitivity: insufficient sample size, uncontrolled variability, and too small an increment of the treatment effect. Recalculation of minimum sample sizes necessary to have both confidence that type I ($p < .05$) and type II ($p > .80$) errors had not been made yielded an estimated sample size of from 100 to 750. These sample sizes are impracticably large and suggest that the issues of variability and the increment of the treatment effect be looked at more carefully.

Examination of variability, as measured by standard deviations, did not show any obvious differences in variability across word-length. This is unsurprising and indeed the purpose of using a within subjects design was to avoid this potential confound. Thus only treatment effects remain as a viable contender for why the word-length effect has thus far been absent.

In examining treatment levels in previous studies in which word-length effects have been present, it was apparent that researchers have often used treatment levels of 1 and 3-syllables (e.g., Caplan, et al., 1992; La Pointe & Engle, 1990) or 1 and 5-syllable words (Baddeley, et al., 1975). Consequently, it was decided to increase the increment in syllable length from a comparison of 1 and 2-syllable words, as was done in Experiments 2 and 4, to a comparison of 1 and 3-syllable word lengths in Experiment 6.

In summary, Experiment 6 presented words that were more discriminable (as measured by syllable-length) than the words used in Experiments 2 and 4, in an effort to increase the sensitivity of the design to word-length effects. There were three critical effects necessary to finding a word-length

effect from which the Articulatory Rehearsal Process could be inferred:

1.1 There would be greater recall of 1 than of 3-syllable words without articulatory suppression;

1.2 There would be lower recall of all items when the person was using articulatory suppression; and

1.3 There would be no, or reduced, word-length effects with concurrent articulatory suppression.

The impact of viewing time on the word-length effect

A second reason why word-length effects may have been absent in Experiments 2 and 4 was because participants spent too much time at input rehearsing the items and hence were retrieving them from a longer term memory instead of working memory, that is, because presentation was participant-paced (W. Parr, personal communication, June, 1993). This explanation is also evidenced by the rather large variability that participants showed in the time spent viewing each stimulus, at times as long as 7 to 10 seconds per item in Experiment 4. It would appear, from a review of the literature, that this question of presentation rate had not been explicitly tested previously in regard to the operationalisation of working memory. The most obvious way to test this hypothesis was to have a control condition in which stimuli were presented at a rate at which word-length effects had been found previously. In the present experiment, Participant-paced and Experimenter-paced conditions were compared to test this hypothesis. The stimulus presentation time was 1.0 second, a time empirically demonstrated as producing word-length effects previously (e.g., Baddeley, et al., 1975). If participant-pacing allowed longer term storage (or some other than working memory process) to be used, then word-length effects would be absent in the Participant-paced condition, but present in the Experimenter-paced condition.

Thus, the fourth hypothesis in Experiment 6 addressed whether an absence of word-length effects in Experiments 2 and 4 was due to persons controlling their own stimulus presentation rate. In Experiment 6 it was predicted that, if participant-pacing allowed use of other than working memory, then:

- 1.4 There would be word-length effects in the Experimenter-paced but not in the Participant-paced conditions.

While Hypotheses 1.1 to 1.4 on their own cannot confirm Postulate 1 (p. 85), support for these hypotheses in a simple-span task is necessary before they are tested in a complex-span task.

Development of measures of content, order, and processing time

In addition to refining experimental procedures, Experiment 6 continued to gather empirical evidence about content, order, and processing time dimensions of working memory.

Content

A major aim of the present research was to examine to what extent what is recalled was dependent on word length and articulatory suppression. In Experiments 2 and 4, wherein word-length was of 1 or 2 syllables, recall was shown to be dependent on whether a person was using articulatory suppression or not, but was independent of word length. Experiment 6 aimed to replicate this dependence on articulatory suppression. Also, it was hoped to examine whether recall depended on word length using more discriminable stimuli.

Finally, Baddeley, et al. (1975) demonstrated that the word-length effect depends more on the time to articulate the words than on syllable length, an effect that is not found with phonological similarity effects (Experiment 5; Schweikert, et al., 1990). For example, when Baddeley, et al.

(1975) presented words of equal syllable lengths, but that took different times to articulate, recall was poorer for words which took longer to articulate. The inference from this result was that the Articulatory Rehearsal Process operated on articulatory features of words. Neither Experiment 2 nor 4 specifically examined how long the stimuli took to articulate. Thus, it is possible that the absence of word-length effects was due to the stimuli not having sufficiently different articulation times. Experiment 6 calculated articulation times after all the stimuli had been presented, to further check that the two word-lengths had different articulation times. Therefore, in Experiment 6, it was predicted that:

1.5 Articulation rate would be predictive of the word-length effect.

Order

In Experiments 1 to 4, the number of words recalled in the correct serial position has averaged 83 % (SD = 6%) of the number of words recalled in any serial position (range 71% - 92%). From these data, it appears that the number of order errors have varied only as a function of phonological similarity; order errors have been unaffected by word-length, articulatory suppression, and sentence verification. From previous experiments in the present dissertation, it can be speculated that the word-length and phonological similarity effects can be differentiated between on the basis that the word-length effect is not produced by order errors but that the phonological similarity is. However, to date, no clear word-length effect has been demonstrated in the present research against which to test this hypothesis directly. Consequently, Experiment 6 aimed to examine the relationship between word-length and order errors explicitly.

If articulatory suppression blocks the Articulatory Rehearsal Process but has no impact on order errors, as suggested above, it is logical to infer that the Articulatory Rehearsal

Process is not directly involved in maintaining order information. That is, it can be hypothesized that:

2.1 Neither articulatory suppression nor word-length effects would affect the number of order errors in each condition.

Processing time

From Experiments 2 and 4 in this dissertation, a distinction between the word-length effect and between processing time at input has been hypothesized. It has been suggested that the word-length effect is an effect produced at output (Avons, et al., 1994; Caplan, et al., 1992; Caplan & Waters, 1994; Coltheart, et al., 1990; Cowan, 1992; Cowan, et al., 1992; Henry, 1991; Monsell, 1987; c.f., Baddeley, & Andrade, 1994). If the word-length effect is not produced at input, then one would not expect to find a relationship between the time spent at input (viewing time) and between the word-length effect. As with Experiment 5, this suggestion was tested in Experiment 6 by hypothesizing that if processing time at input has no effect on maintenance or output processes then:

2.2 Viewing time would be unrelated to changes in recall due to word-length and articulatory suppression.

Meta-memory

Finally, the third main purpose of Experiment 6 was to examine the accuracy of persons' on-line meta-memory. In Experiment 5, data indicated that persons were unable to predict the phonological similarity effect, an effect hypothesized as indicating the operation of the Phonological Store of working memory. Experiment 6 examined whether persons could predict the word-length effect and, by implication, the Articulatory Rehearsal Process. In Experiment 6, as in Experiment 5, pre-estimates were made with participants being aware of the word-length condition and whether they will be using articulatory suppression or

not. It was anticipated that these informed pre-estimates would be highly correlated to recall performance.

Summary

To summarize, the purpose of Experiment 6 was to overcome methodological shortcomings due to inadequate sample size and low stimulus discriminability, and to compare the effect of participant versus experimenter-paced stimulus presentations on word-length effect. Second, Experiment 6 aimed to further develop the validity of using measures of content, order, and processing time when a word-length effect was present. Finally, Experiment 6 aimed to examine how accurate on-line meta-memory is when there is also a word-length effect in objective performance. Together, these three aims were intended to lay the empirical basis for replication of the word-length effect, and for demonstrating the effect of the Articulatory Rehearsal Process in a quantitative/process model of working memory.

Method

Participants

Participants in Experiment 6 were the same participants as in Experiment 5. There were 34 people in Experiment 6; 17 in the Experimenter-paced condition (Mean age = 24.35, SD = 8.71) and 17 in the Participant-paced condition (Mean age = 21.71, SD = 2.64). There was no age difference between pacing conditions, $t(32) = 1.20$, $p = .245$, two-tailed. Due to a computer error, pre-estimate data for one participant was lost, reducing the sample size for that variable from $N = 34$ to $N = 33$.

Design

The design was a three-factor mixed design. The factors of word-type (1-syllable, 3-syllable) and articulatory

suppression (control, articulatory suppression) were varied within subjects. The factor of pacing (Experimenter-paced, Participant-paced) was varied between subjects. The dependent variables were as detailed in the General Method section (p. 99).

Procedure

Prior to arrival, participants were randomly placed in either an Experimenter-paced, or Participant-paced condition. Participants in the Experimenter-paced condition were presented with words at the rate of one per second. Participants in the Participant-paced condition were presented with their next target word when they pressed the spacebar. Otherwise the procedure has been described fully in the General Method section (p. 95).

Results

The following results of Experiment 6 showed that the word-length effect was replicated when stimuli were discriminable by more than 1 syllable in length. Measures of content, order, and processing time continued to be useful in defining specific features of the word-length effect. Finally, persons were accurate in predicting both direction and magnitude of the separate, but not interactive, contributions of articulatory suppression and word-length.

Content Measures: Recall in the correct serial position and recall in any serial position

Recall in the correct serial position

Overall there was no difference between Experimenter-paced and Participant-paced presentations for recall in the correct serial position. As predicted, there was lower recall of 3- than of 1-syllable words (in the control condition), $F(1,32) = 9.00$, $p = .005$, $MSE = .42$, confirming Hypothesis 1.1. There was an overall decline in recall with articulatory

suppression, $F(1,32) = 134.90$, $p < .001$, $MSE = .46$, confirming Hypothesis 1.2. There was also a statistical trend toward an interaction between word-length and articulatory suppression, $F(1,32) = 3.79$, $p = .06$, $MSE = .35$, partially supporting Hypothesis 1.3. The trend-level interaction between word-length and articulatory suppression occurred because in the control condition there were fewer 3-syllable words recalled than 1-syllable words but with articulatory suppression there was no difference in recall between 1 and 3-syllable words, as would be predicted by Baddeley, et al.'s (1975) description of the word-length effect. Finally, there were no significant differences between post-estimates of recall in the correct serial position and actual performance when compared using a paired t-test (Table 30). This result indicated that participants were probably not guessing item positions.

Table 30. Mean (SD) recall of 1-syllable or 3-syllable words either in the correct serial position or in any serial position for both the control and articulatory suppression conditions across experimenter and participant-paced presentations in Experiment 6 ($N = 34$).

	Mean Recall (SD)		
	Recall in the correct serial position	Recall any serial position	Post-estimates of recall in the correct serial position
Experimenter paced			
Control			
1-syllable	3.95 (1.15)	4.49 (0.89)	3.98 (1.05)
3-syllable	3.39 (1.24)	4.16 (0.88)	3.43 (1.11)
Articulatory suppression			
1-syllable	2.36 (0.90)	3.04 (0.88)	2.50 (0.62)
3-syllable	2.25 (0.69)	3.01 (0.79)	2.23 (0.99)
Participant paced			
Control			
1-syllable	4.41 (0.96)	4.88 (0.74)	4.49 (0.62)
3-syllable	4.03 (1.03)	4.56 (0.95)	4.16 (1.00)
Articulatory suppression			
1-syllable	2.90 (0.99)	3.52 (0.73)	3.00 (0.94)
3-syllable	2.87 (0.79)	3.45 (0.58)	2.90 (0.79)

Thus, data in Experiment 6 show a word-length effect. Furthermore, contrary to Hypothesis 1.4, whether words were presented at a fixed rate or self-paced made no difference to recall in the correct serial position. Irrespective of presentation rate, there was a word-length effect without articulatory suppression, which was abolished by articulatory suppression (Table 30; Figure 13).

Recall in any serial position

For recall in any serial position the effects were less robust but paralleled those of recall in the correct serial position. In the control condition, one-syllable words were better recalled than three-syllable words, $F(1,32) = 3.14$, $p = .096$, $MSE = .19$ only at a statistical trend level (Hypothesis 1.1). Articulatory suppression lowered recall overall, $F(1,32) = 246.55$, $p < .001$, $MSE = .22$ (Hypothesis 1.2). As predicted by Hypothesis 1.3, there was no word-length effect with articulatory suppression, although the implied interaction between word-length in the control and articulatory suppression conditions again failed to reach significance, $F(1,32) = 4.06$, $p = .052$, $MSE = .15$, at other than a statistical trend level. Finally, there was no overall effect on recall whether stimuli were presented at an experimenter or at a participant-determined pace (Hypothesis 1.4). In summary, recall in any serial position did not show a word-length effect except at the statistical trend level (Table 30).

In Experiment 5, it was found that word-type interacted with whether order information was required or not. So, in Experiment 6, the parallel question was raised of whether there was a difference in word-type, in this instance the word-length effect, when order information was not required. Statistically this would appear as an interaction between recall method and word-length, an interaction that failed to reach significance in Experiment 6, $F(1,33) = 2.09$, $p = .158$, $MSE = .06$. Thus, there was no basis to link the word-length

effect with order information in Experiment 6. This is an important result because it implies that the word-length effect is an effect of content and not of order; a conclusion in line with the original description of the Articulatory Rehearsal Process (Baddeley, 1986, 1992a, 1992b). Conversely, this result also implies that order information is not acted upon by the Articulatory Rehearsal Process, with the implication that one can sensibly dissociate the Phonological Store, which is concerned with order information, from the Articulatory Rehearsal Process which is not. This conclusion will be amplified in the discussion on order below.

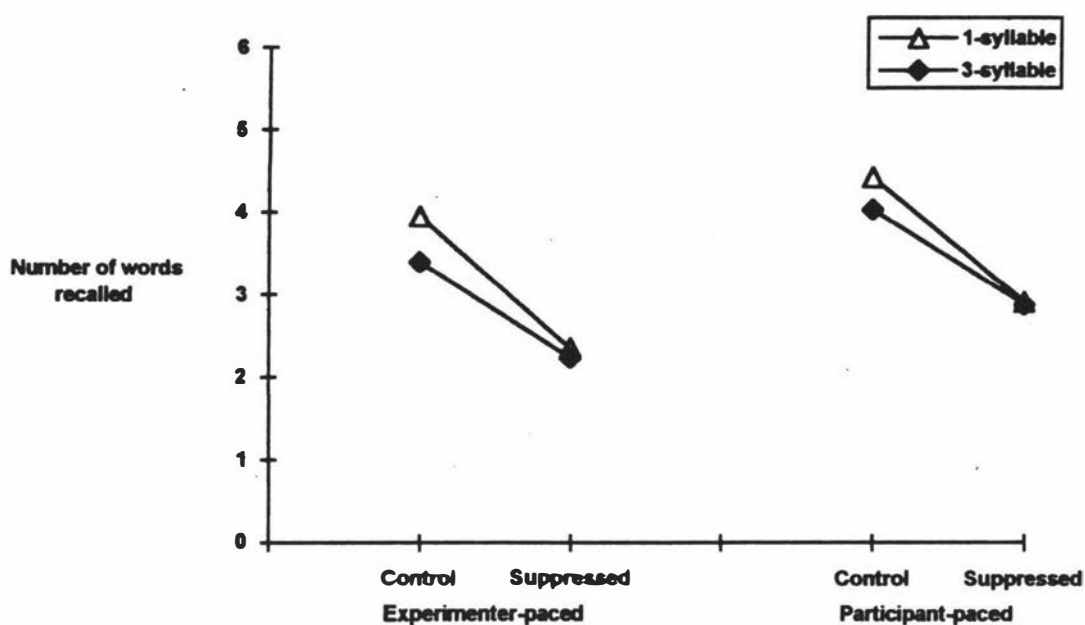


Figure 13. Recall in the correct serial position by word length, articulatory suppression, and pace in Experiment 6.

Articulation rate and its effect on word-length

There was a slower articulation rate for 3-syllable than for 1-syllable words, $F(1,32) = 135.19$, $p < .001$, $MSE = .05$, irrespective of any other factors. There were no other main nor interaction effects (Table 31). This slower rate of articulation replicated the fifth experiment of Baddeley, et al. (1975). Hypothesis 1.5 predicted that articulation rate would be predictive of recall. To test this hypothesis, measures of content were regressed against measures of

articulation rate. Both variables were again transformed into their natural logarithms to reduce non-linearity of regression coefficients (Keppel, 1982).

Table 31. Articulation rate (SD) for 1-syllable or 3-syllable words for both the control and articulatory suppression conditions and across pacing conditions in Experiment 6 (N = 34).

	Experimenter-paced		Participant-paced	
	Control	Suppressed	Control	Suppressed
1-syllable	2.03 (0.44)	2.02 (0.36)	1.96 (0.59)	1.98 (0.54)
3-syllable	1.68 (0.26)	1.70 (0.28)	1.62 (0.37)	1.67 (0.40)

Unlike in Experiment 5, in Experiment 6 regressing storage measures against articulation rate produced disappointing results; no regression coefficients reached significance. Thus, although there were differences in articulation rate as a function of word-type, these differences did not appear to be predictive of the word-length effect per se.

Order: The impact of the word-length effect upon order errors

Overall there were no significant differences in order errors across any conditions and no interactions between conditions (Hypothesis 2.1; Table 32). That is, word-length appeared to have no effect upon order errors.

Table 32. Mean (SD) order errors for 1-syllable and 3-syllable words for both the control and articulatory suppression conditions and across the pacing conditions in Experiment 6 (N = 34).

	Order errors	
	Experimenter-paced (N = 17)	Participant-paced (N = 17)
Control		
1-syllable	0.54 (0.46)	0.47 (0.40)
3-syllable	0.76 (0.58)	0.56 (0.46)
Articulatory suppression		
1-syllable	0.68 (0.51)	0.62 (0.42)
3-syllable	0.77 (0.41)	0.58 (0.44)

Finally, in order to make the analysis of order information as complete as was done in Experiment 5, an analysis of the

item transpositions was done for data in Experiment 6. As can be seen in Figure 14, item transpositions were typically close to the correct serial position and more item transpositions occurred in the middle than the outer serial positions. These results are equivalent to the results of Experiment 5. However, unlike Experiment 5, the data from Experiment 6 showed that there were no more item transpositions for 1-syllable than for 3-syllable words using a Chi-squared test at $p < .05$. This result is consistent with the statement that the word-length effect, and by inference the Articulatory Rehearsal Process, has no effect on item transposition.

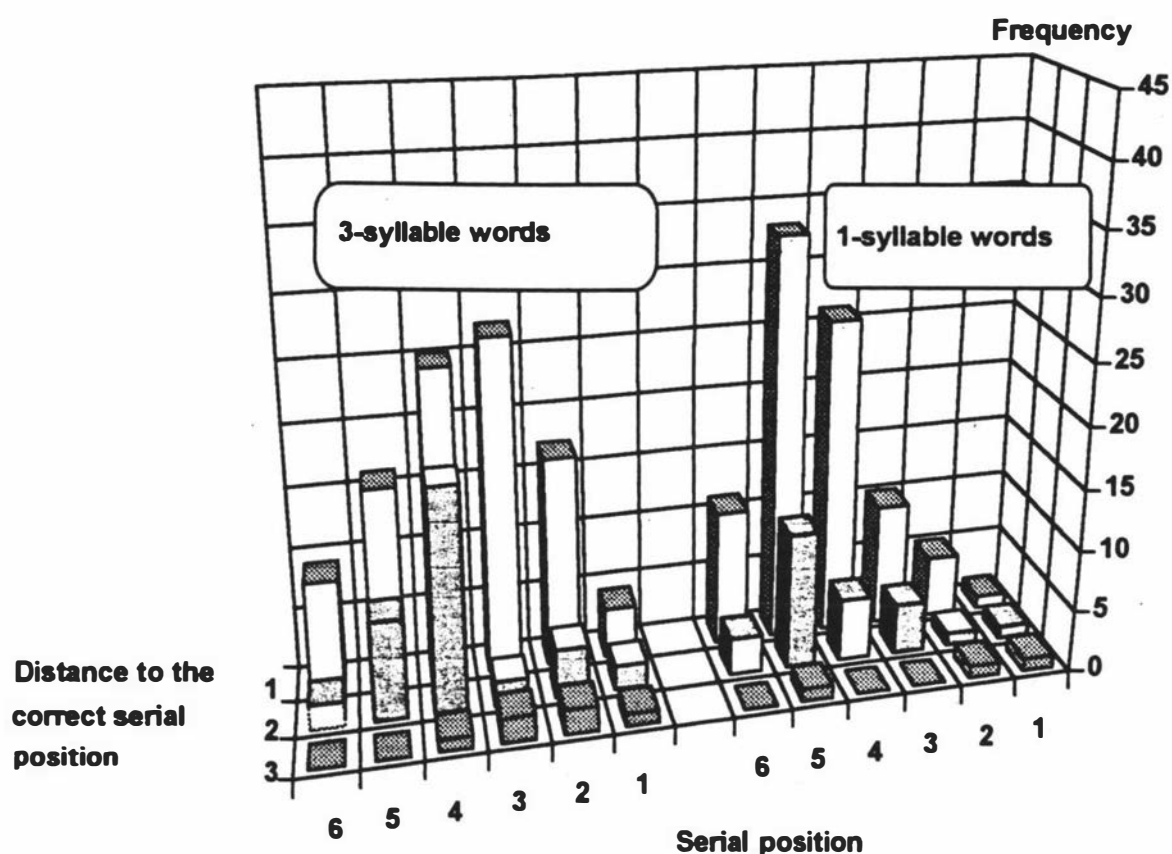


Figure 14. Item transpositions for 1- and 3-syllable words in Experiment 6.

In summary, the results from Experiment 6 showed that, while the length of an item may create some differential loss of how many items can be recalled, word-length has no impact upon recall of order information. That is, if one assumes that the word-length effect is produced by the Articulatory

Rehearsal Process, then one might also conclude from the present data that the Articulatory Rehearsal Process is not involved with the maintenance of order information.

Processing time: The impact of viewing time on the word-length effect

Overall, as might be expected, persons who were able to self-pace presentation of stimuli spent longer viewing each item than persons who were presented at a fixed rate of 1 word per second, $F(1,32) = 80.36, p < .001, MSE = 758,739$. For the Participant-paced condition, the range of means across word-length and articulatory suppression conditions was from 2,278 to 2,403 ms (SD = 662 to 726 ms). There were no differences in participant-pacing across either word-length nor articulatory suppression conditions (Table 33).

It was suggested earlier that participant-pacing might affect the word-length effect if the extra time allowed processes other than the Phonological Loop to operate. A regression of the two content measures onto mean viewing time failed to show any relationship between the input process and later recall. This important result again suggests the conclusion that the Articulatory Rehearsal Process is unaffected by the processes operating at input of the stimuli. That is, time spent viewing each item does not vary as a function of either word-length nor articulatory suppression.

Table 33. Mean (SD) viewing time (ms) for both 1-syllable and 3-syllable words for both the control and articulatory suppression conditions and across pacing conditions in Experiment 6 (N = 34).

	Viewing times	
	Experimenter-paced (N = 17)	Participant-paced (N = 17)
Control		
1-syllable	1000	2340 (726)
3-syllable	1000	2336 (685)
Articulatory suppression		
1-syllable	1000	2278 (662)
3-syllable	1000	2403 (696)

On-line meta-memory: Pre-estimates of recall in any serial position

Overall pre-estimates of recall in any serial position were higher for the Participant-paced than the Experimenter-paced condition, $F(1,31) = 5.87$, $p = .021$, $MSE = 1.37$ and pre-estimates were lower with articulatory suppression, $F(1,31) = 103.60$, $p < .001$, $MSE = .34$. Pre-estimates were higher for 1- than 3-syllable words in both the control condition, $F(1,31) = 71.49$, $p < .001$, $MSE = .10$, and the articulatory suppression condition, $F(1,31) = 23.92$, $p < .001$, $MSE = .11$ (Table 34). Thus, persons appear able to monitor their working memory sufficiently well to be sensitive to the impact of word-length and of articulatory suppression separately. However, persons were unable to predict that articulatory suppression would remove the word-length effect, an interaction effect. From these results, it is concluded that participants are able to predict the pattern of main effects that occur when stimuli are altered across word-length and articulatory suppression.

Table 34. Mean (SD) pre-estimates of recall in any serial position for both 1-syllable and 3-syllable words for both the control and articulatory suppression conditions and across pacing conditions in Experiment 6 ($N = 33$).

	Pre-estimates	
	Experimenter-paced ($N = 17$)	Participant-paced ($N = 17$)
Control		
1-syllable	4.05 (0.66)	4.37 (0.91)
3-syllable	3.25 (0.62)	3.83 (0.70)
Articulatory suppression		
1-syllable	2.76 (0.42)	3.34 (0.74)
3- syllable	2.41 (0.77)	2.89 (0.61)

Having established that persons could predict the pattern of word-length effects, the next step was to ascertain whether there was a regular relationship between predictions and performance. Because there were no differences across the pacing conditions, the data were collapsed across pacing condition to examine the pre-estimate-performance

relationship. Regressing recall in any serial position against pre-estimates showed that, in the control condition, there was a relationship between pre-estimates and performance for 1-syllable, $R^2 = .30$, $F(1,31) = 13.75$, $MS_{\text{residual}} = 0.49$, and 3-syllable words, $R^2 = .19$, $F(1,31) = 7.74$, $MS_{\text{residual}} = 0.01$ at $p < .05$. There was also a statistically significant relationship at $p < .05$ between pre-estimates and performance measures for 1-syllable and for 3-syllable words in the articulatory suppression condition, $R^2 = .18$, $F(1,31) = 6.87$, $MS_{\text{residual}} = 0.01$, and $R^2 = .19$, $F(1,31) = 7.20$, $MS_{\text{residual}} = 0.01$, respectively. That is, in addition to predicting the pattern of their performance, there was a significant relationship in each condition between pre-estimates and performance. From these results one can conclude that persons demonstrated accurate on-line meta-memory for the word-length effect.

Discussion

The present study examined four aspects of the word-length effect and working memory: the parameters for replication of the word-length effect; the relationship of articulation rate to recall; the use of measures of content, order, and processing time; and the effectiveness of on-line meta-memory.

The parameters necessary for replication of the word-length effect

There were two possible hypotheses why the word-length effect was absent in Experiments 2 and 4. First, it was hypothesized that the participant-paced stimulus presentation mode allowed some type of longer term storage to occur, a form of storage which did not produce a word-length effect. Second, it was hypothesized that the stimuli used were not sufficiently discriminable to produce a word-length effect.

The data from Experiment 6 clearly showed that the presentation mode did not affect the magnitude of the word-

length effect. From this result it is concluded that participant pacing is not likely to have been the reason that word-length effects were not replicated in Experiments 2 and 4. Furthermore, this result also implies that participant pacing, at least within the parameters used in the present studies, probably does not allow persons to encode information into a longer term memory in preference to working memory.

In contrast, given that Experiment 6 replicated the word-length effect, it would appear that the conclusion that the previous failures to replicate the word-length effect in the present research were due to stimuli being insufficiently discriminable. The basis for this conclusion is two-fold. First, in a meta-analysis of Experiments 2 and 4 a slight but significant word-length effect was found. Thus, an increase in the degrees-of-freedom was sufficient to find a word-length effect. Second, in Experiment 6, discriminability was increased by altering the size of the treatment effect (increasing the contrast to 1 and 3-syllables) which produced a reliable word-length effect.

In conclusion, it would appear that the word-length effect requires stimuli different by at least 2-syllables in length. While, in retrospect, this result is perhaps unremarkable, at the outset of the present dissertation a review of the literature suggested that, provided there was a difference in the time to articulate the stimuli, a word-length effect would be present (Baddeley, 1986, 1992a, 1992b). Experiment 6 (when contrasted with Experiments 2 and 4) demonstrates that this inference applies less universally than was first thought. This replication of the word-length effect also raises the question of whether there was a predictive difference in articulation rate that might have produced that effect.

The relationship of articulation rate to recall

Experiment 6 also aimed to replicate the work of Baddeley, et al. (1975) in which word-length effects were shown to be related to articulation rates of the stimulus words. In the present study, it was expected that articulation rate would be predictive of the word-length effect. However, there was no predictive relationship shown between articulation rate and recall, replicating Tehan and Humphry's (1988) results. The implication of this result is that the present data fail to support a model of the word-length effect being directly produced by differing rates of articulatory-based rehearsal.

The use of measures of order and processing time

Of most relevance to further defining the word-length effect was the result that requiring the order of items did not differentially affect the word-length effect.

It was clear that, in contrast the phonological similarity effect, the word-length effect is in no way influenced by order information. There was a constant number of order errors irrespective of word-length or articulatory suppression. Using the more detailed of analysis of transposition errors (from Healy, 1974, 1982) confirmed that order information was unaffected by word-length. Taken together, the findings of no differential effect of requiring order to be recalled, no difference in order errors, and no difference in transposition errors strongly support the conclusion that the word-length effect is not directly involved with order information is warranted.

Finally, the measure of processing time contributed very little within-subject information to either a measure of content or of order. This null result is, however, important as it provides further evidence that input processes are separate from output processes (Avons, et al., 1994; Caplan, et al., 1992; Caplan & Waters, 1994; Coltheart, et al., 1990;

Cowan, 1992; Cowan, et al., 1992; Henry, 1991; Monsell, 1987; c.f., Baddeley, & Andrade, 1994).

The effectiveness of on-line meta-memory

The third purpose of Experiment 6 was to investigate how accurately persons were able to predict their subsequent recall. Experiment 6 extends previous experiments by showing that persons can accurately predict the word-length effect and the effect of articulatory suppression on recall, but not the removal of word-length effects with articulatory suppression. The present results suggest that persons can monitor the part of the working memory that produces a word-length effect -- the Articulatory Rehearsal Process (Baddeley, 1986, 1992a, 1992b; Baddeley, et al., 1975).

SUMMARY OF EXPERIMENTS 5 AND 6

Experiments 5 and 6 have collectively provided some valuable information. The parameters necessary for replicating the phonological similarity and word-length effects, respectively, have been delineated. Order information has been shown to be critical to the phonological similarity but not the word-length effect, and the role of processing time has been shown to be uninfluential on subsequent recall within the present parameters. Finally, it would appear that persons can monitor the word-length but not the phonological similarity effect.

Experiment 5 explicates the parameters for replicating the phonological similarity effect: stimuli from a fixed pool and with limited recall time. The phonological similarity effect appears to occur because of a greater loss of order information for words that are phonologically similar over words that are not. When confusability was increased by not only presenting phonologically similar words, but by repeating those words in different orders over trials, the phonological similarity effect was exacerbated. Also, the

present data suggest that processing time at input, within certain limits, does not alter subsequent recall.

Experiment 6 shows that replication of the word-length effect requires that treatment effects be discriminable by at least two syllables. Also, whether stimuli were presented at an experimenter or participant-paced rate did not influence the magnitude of the word-length effect. Taken as a package, the results in Experiment 6 relating to content, order, and processing time raise some disquiet about the word-length effect. Specifically, the word-length effect appears to have been produced without a causal relationship to articulation rate. Second, the word-length effect, supposedly produced because of longer words taking more rehearsal time, seemed to be unrelated to the time taken at stimulus input, a time when supposedly one would be rehearsing.

However, Experiments 5 and 6 have assisted in the integration of working memory into a coherent model in three ways. First, the conditions necessary for replicating the phonological similarity and word-length effects have been specified more clearly (Figure 3. p.87). This is important because these effects have been used to define the Phonological Store and the Articulatory Rehearsal Process, respectively, of the qualitative/structural model of working memory (Baddeley, 1986, 1992a, 1992b). In replicating the phonological similarity and word-length effects, there has also been some valuable information gathered about the nature of these effects that will be used in the General Discussion section to critique some of the inferences made from these effects. Fundamentally, the way is now open to explicitly compare how the phonological similarity and word-length effects operate in complex-span versus simple-span tasks. That is, the prerequisite steps have been taken for a direct test of Postulate 1.

Second, Experiments 5 and 6 again demonstrate the utility of using measures of content, order, and of processing time as

differential indicators of the phonological similarity and word-length effects. That is, Experiments 5 and 6 have shown that qualitative/structural effects of working memory have different effects on quantitative/process variables of working memory, and thus support Postulate 2.

Third, Experiments 5 and 6 provide information about the ability of persons to monitor qualitative/structural aspects of their working memory (Postulate 3). It was suggested from these data that persons can monitor the Articulatory Rehearsal Process (as indicated by the word-length effect) but not the Phonological Loop (as indicated by the phonological similarity effect). The implication of this finding is that maybe only some aspects of working memory can be monitored. This point will be expanded on further in the General Discussion section.

In conclusion, Experiments 5 and 6 represent a critical step in preparing for a direct test of the assertion that the qualitative/structural and quantitative/process models of working memory measure and describe the same construct. In providing a theoretical framework for the phonological similarity and word-length effects, Experiments 5 and 6 lead directly to Experiment 7 in which those effects are compared across simple and complex-span tasks.

CHAPTER 10: EXPERIMENT 7: THE INTEGRATION OF A QUALITATIVE/STRUCTURAL MODEL OF WORKING MEMORY WITH A QUANTITATIVE/PROCESS MODEL OF WORKING MEMORY

To date, the present research has been examining some steps necessary before directly addressing the following question: *Are the two models of working memory in fact examining the same construct?* The present research has also been examining the question of whether *persons can monitor their working memory*. In addressing the first question, to date evidence has been found from Experiments 1 to 6 that shows a need for dimensions of order and processing time, in addition to content, to be considered when describing the qualitative/structural model of working memory (Postulate 2, p.88). Also, in addressing the second question, Experiments 1 to 6 have shown that persons can monitor their working memory in both a general way and can also specifically monitor their Articulatory Rehearsal Process as indicated by the word-length effect (Postulate 3, p.88). What remains to be done with this present research series is to (1) directly investigate the presence of phonological similarity and word-length effects in both simple-span and complex-span tasks (Postulate 1, p.85), to (2) investigate whether the magnitude of those effects alters as a function of whether a concurrent operation is present or not, and to (3) replicate previous findings as regards content, order, processing time, and meta-memory.

As regards the first aim of Experiment 7, only an indirect comparison has been made between simple-span and complex-span tasks in the present research (Table 35). Experiment 7 directly compared simple and complex-span tasks in a single experiment (Figure 3, p. 87). In comparing the simple and complex-span tasks directly, the hypothesis was that if the two models refer to the same construct of working memory, then qualitative effects (word-length and phonological

similarity) would be equally apparent in both models (Postulate 1). Furthermore, Experiment 7 also investigated whether the magnitude of those effects altered when a concurrent task was present or not. The importance of these first two aims lies in demonstrating that those effects originally used to infer the Phonological Loop (phonological similarity and word-length effects; Baddeley, 1986, 1992a, 1992b) also occur in a quantitative/process model of working memory. The importance of demonstrating co-occurrence of effects and dimensions is that if both models show comparable effects, then one cannot posit that they are theoretically separate accounts of working memory, as has been done (Gathercole & Baddeley, 1993; Just & Carpenter, 1992). However, to date, only data supporting Postulate 2 has been found. To provide some convincing evidence that both models of working memory refer to the same construct there needs to be data supporting Postulate 1 and 2 concurrently. Thus, the primary aim of Experiment 7 was to investigate whether phonological similarity and word-length effects co-occur in simple and complex-span tasks.

Table 35. A summary of the domains of previous experiments in the present research.

	Task-type	
	Simple-span	Complex-span
Qualitative effect		
Phonological similarity	Expts. 3 & 5	Expt. 1
Word-length	Expts. 4 & 6	Expt. 2
Articulatory suppression	Expts. 3, 4, 5, & 6	Expts. 1 & 2

In addition, Experiment 7 aimed to replicate the use of content, order, and processing time as three necessary dimensions of working memory in both simple and complex-span tasks (Postulate 2). In this series of experiments, these dimensions have only been investigated in a simple-span task when phonological similarity (Experiment 5) and word-length (Experiment 6) effects were present.

Finally, in this series of experiments it has been suggested that persons could not monitor the phonological similarity

effect, and by inference could not monitor the operation of the Phonological Store (Experiment 5). In contrast, it was shown that persons were able to monitor the word-length effect and by inference the Articulatory Rehearsal Process (Experiment 6). Experiment 7 aimed to replicate these findings prior to drawing firmer conclusions about persons' ability to monitor aspects of their working memory.

In summary, this research aimed to integrate two models of working memory via a direct comparison of those models. Demonstrating the unity of the two models of working memory was also expected to allow the development of a single more complete model of working memory using dimensions of content, order and processing time; this was a second goal of the present research program. The third goal of this research was to examine the manner and degree to which this more complete model of working memory is available for on-line meta-memorial monitoring.

Differences between simple and complex-span tasks: A review of literature to date

The literature directly examining the common features of the qualitative/structural and quantitative/process definitions of working memory is exceedingly sparse -- one study in fact. La Pointe and Engle (1990) compared simple and complex-span tasks with respect to the word-length effect. As reviewed earlier (pp. 76-78), La Pointe and Engle found that *both* simple and complex-span tasks showed a word-length effect and that recall was generally higher in the simple-span task than in the complex-span task. This result is not unreasonable given that a complex-span task, by its very definition, probably involves more processing in the same time period than a simple-span task. La Pointe and Engle also found that the word-length effect was greater for simple than for complex-span tasks. This interaction between word-length and

task-type is more interesting. La Pointe and Engle commented that

It is not clear that the task [complex-span or simple-span] by word length interaction has any interesting implications. ...it is likely the case that task main effects and interactions involving task occur because the complex-span tasks are more constrained by performance boundaries than the simple-span tasks. " (p.1121).

However, if one accepts Baddeley's (1986, 1992a, 1992b) account of the word-length effect, an alternative interpretation is possible. A smaller word-length effect in the complex-span task could suggest that less rehearsal by the Articulatory Rehearsal Process had occurred in the complex-span task than in the simple-span task. Far from being unimportant, the implication of there being less rehearsal in a complex-span is that the sentence verification process competes for the same resource as the Articulatory Rehearsal Process.

Finally, there are some methodological issues that, while not critical to the results of La Pointe and Engle (1990), certainly deserve comment and empirical investigation. First, their complex-span task was unlike a more typical Reading Span task in that the target words were not part of the sentence but were simply at the end of the sentence. This contrasts with the present Experiment 7 which presented target words that were part of the sentence being verified. Second, La Pointe and Engle's complex-span task was participant-paced, where their simple-span task was experimenter-paced. While in Experiment 6 of the current research it was suggested that type of pacing does not significantly alter the word-length effect, Experiment 7 nonetheless presented both simple and complex-span tasks as self-paced to allow direct examination of viewing time differences and to thus examine any speed-accuracy tradeoffs.

Third, La Pointe and Engle used an incremental span task, typical of variants of the Reading Span task. Experiment 7, in contrast, used a fixed-span presentation of stimuli. It is unclear at this point if this difference has any theoretical importance above the pragmatic point that fixed-span tasks allow calculation of order errors -- a score critical to the assessment of order in the present experiments.

The Present Study

Having determined that both the qualitative/structural and quantitative/process models can be described in process terms, having identified parameters necessary to production of the phonological similarity and the word-length effects, and having reviewed research examining the integration of simple and complex-span tasks, the present research returned to the question of whether both definitions of working memory reflect the same construct. That is, are there differences between simple-span and complex-span operationalisations of working memory?

From La Pointe and Engle (1990), the implication is that there are quantitative differences in favour of simple-span tasks with recall. In contrast, a meta-analysis between Experiments 1 & 2 (complex-span) and Experiments 3 & 4 (simple-span) suggested that there were no quantitative differences between simple-span and complex-span tasks. Notwithstanding this difference across studies of whether there are quantitative differences in the word-length effect or not, both La Pointe and Engle and the present studies suggest that there are no qualitative differences between simple-span and complex-span tasks (in agreement with Postulate 1). That is, it could be expected that the word-length effect would occur in both simple-span and complex-span tasks.

In addition, if the word-length effect occurs in both simple-span and complex-span tasks, an effect that is indicative of

the Articulatory Rehearsal Process of the Phonological Loop, then it is not unreasonable to infer that the phonological similarity effect, an effect indicative of the phonological Store of the Phonological Loop would also occur in both task-types. Consequently, for Experiment 7, the main hypothesis was that *there would be no qualitative differences between simple-span and complex-span tasks across the three dimensions of content, order, and processing time*. That is, the presence of the phonological similarity and word-length effects will be unaffected by whether target items are presented in a simple or in a complex-span task. This main hypothesis can be divided into more specific hypotheses based on the three postulates this research is investigating.

Hypotheses

Hypotheses for Experiment 7 will now be presented in terms of how they impact upon Postulates 1 to 3.

Hypotheses relating to Postulate 1

The primary aim of Experiment 7 was to demonstrate equivalent phonological similarity and word-length effects in both simple and complex tasks. Thus, the following four hypotheses were proposed :

- 1.1 There would be a phonological similarity effect in both simple-span and complex-span tasks,
- 1.2 The phonological similarity effect would not be different between simple-span and complex-span tasks,
- 1.3 There would be a word-length effect in both simple-span and complex-span tasks, and
- 1.4 The word-length effect would not be different between simple and complex tasks.

In Experiments 5 and 6, articulation rate showed an equivocal relationship to recall. So, Experiment 7 re-examined the predictiveness of phonological similarity and word-length effects from articulation rate.

Hypotheses relating to Postulate 2

From previous experiments in the current research, especially Experiments 5 and 6, it was expected that:

2.1 Order errors would be greater for phonologically similar words, and

2.2 Order errors would be invariant across word-length.

Additionally, it was investigated whether task-type would have any effect on order errors, either in magnitude or direction.

Processing time has been an especially elusive concept to define. However, it seems reasonable to make two suppositions. First, the simple-span task will require less total processing time than a task with an additional sentence verification component. Second, the process of working memory is time limited, to about 2 seconds (Baddeley, 1986). So, in a limited amount of time one of two things might occur. The increased processing requirements of the complex over the simple task might result in more time being spent viewing stimuli in the complex task, perhaps to offset some reduced processing efficiency. Alternately, there may be no major loss of information or of speed if the word recall component of working memory, the Phonological Loop, can operate in parallel with the component responsible for sentence verification, the Central Executive (Gathercole & Baddeley, 1993; Just & Carpenter, 1992). Data from Experiment 6 showed that increased processing time for participant-paced over experimenter-paced presentations in a simple-span task resulted in no difference in recall. It has also been argued

earlier in the present dissertation that the input process is separate from the maintenance/output process. Also, the original studies by Baddeley and Hitch (1974) showed that a recall operation was not necessarily disrupted by other tasks using what was described as a short term store. So, for Experiment 7 the hypothesis was that:

2.3 Viewing times would be independent of recall.

Hypotheses relating to Postulate 3

From previous experiments in the current research using an on-line meta-memory task, it was expected that:

3.1 persons would not be able to predict the phonological similarity effect,

and that

3.2 persons would be able to predict the word-length effect.

In order to predict either of the above effects, persons were expected to be able to predict the pattern of their performance. Second, it was expected that there would be a predictive relationship of performance by pre-estimates.

Method

Participants

Thirty-two students from a second year psychology class were offered either course credit or a small honorarium for participation in Experiment 7. The mean age for the 24 female students was 23.21 (SD = 7.50) years and for the 8 male students was 22.50 (SD = 2.98) years. There were no significant age differences between the male and female groups.

Design

Experiment 7 was a two-factor within-subjects design. The factors of word-type (control, phonologically similar, and 3-syllable) and task-type (simple-span, complex-span) were varied within subjects. The control condition involved 1-syllable, phonologically dissimilar target words. The dependent variables were as detailed in the General Method section (p. 99).

Procedure

Having established by the end of Experiment 6 that the effects being obtained were robust, and although no practice effects were evident, it was decided to further check for practice effects. So, unlike in Experiments 1 to 6, Experiment 7 used a counterbalanced block presentation across persons. In Experiment 7 the order of each group of 3 trials for each word-type condition (control, phonologically similar, and 3-syllable) was counterbalanced, giving a total of 9 trials and forming one block. The order of the resulting block was then reversed for the second block of trials. The order of the first two blocks was then repeated for the third and fourth blocks, giving a total of four blocks and 36 trials. Finally, whether the first or last two blocks were simple-span or complex-span tasks was counterbalanced across participants.

Also, in Experiment 7, articulatory suppression was not included as a condition. The effect of articulatory suppression has been shown in Experiment 5 and 6 to remove both phonological similarity and word-length effects. Thus, when comparing whether those effects are present in both simple-span and complex-span tasks, it did not seem useful to include a condition that prevented those effects and to then compare whether those effects were prevented equally in simple-span and complex-span tasks. Also, by not having an articulatory suppression condition, a greater number of

experimental trials were possible for the conditions of interest.

Apart from the above difference in presentation order, the procedure for Experiment 7 has been fully described in the General Method section (p. 95) and paralleled that of previous experiments in the present series.

Results

The results section for Experiment 7 will follow the format established in the previous 6 experiments and present results for the phonological similarity and word-length effects consecutively within each section.

Content measures: Recall in the correct serial position and recall in any serial position

The purpose of examining content in Experiment 7 was two-fold: to replicate the phonological similarity and word-length effects; and, most importantly, to directly compare the effect of task-type upon the phonological similarity and word-length effects. The data showed that the phonological similarity and word-length effects were replicated. There was no impact of task-type upon the phonological similarity effect. However, task-type did increase the magnitude of the word-length effect (Table 36; Figure 15).

There were four hypotheses to be addressed regarding measures of content. It was expected that there would be main effects of phonological similarity and of word-length. It was also expected that neither the phonological similarity nor word-length effects would be influenced by whether the task was a simple-span or complex-span task. That is, no interaction was expected between word-type and task-type.

Table 36. Mean (SD) recall of control, phonologically similar, and 3-syllable words either in the correct serial position or in any serial position by task-type (simple or complex) for Experiment 7 ($N = 32$).

	Recall	
	Recall in the correct serial position	Recall in any serial position
Simple-span		
Control	4.02 (1.40)	4.75 (1.05)
Similar	3.10 (1.14)	4.57 (0.97)
3-syllable	3.50 (1.26)	4.51 (0.77)
Complex-span		
Control	3.55 (1.19)	4.44 (0.67)
Similar	2.90 (1.01)	4.16 (0.71)
3-syllable	2.79 (0.93)	3.85 (0.85)

Recall in the correct serial position

As predicted by Hypothesis 1.1, fewer phonologically similar words were recalled than control words, $F(1,31) = 34.32$, $p < .001$, $MSE = .57$. Hypothesis 1.2 predicted that there would be no main effect of task-type, a prediction supported by the data, $F(1,31) = 2.33$, $p = .137$, $MSE = 1.55$, nor an interaction between task-type and word-type, again supported by the data, $F(1,31) = 1.07$, $p = .308$, $MSE = .49$.

Hypothesis 1.3 predicted a word-length effect between the control and 3-syllable word-type conditions. Data showed that fewer 3-syllable words were recalled than the control condition words, $F(1,31) = 32.10$, $p < .001$, $MSE = .41$. Hypothesis 1.4 predicted that there would be no main effect of task-type, nor an interaction between task-type and word-type. The data showed that there were more words recalled in the simple-span than in the complex-span condition, $F(1,31) = 4.76$, $p = .037$, $MSE = 2.31$. However, the magnitude of the word-length effect did not change with this decrease in recall for the complex-span task, as inferred from the absence of an interaction between word-type and task-type, $F(1,31) = 1.38$, $p = .248$, $MSE = .35$ (Table 36; Figure 15). Thus, while overall recall in the correct serial position was

lower in the complex-span task than in the simple-span task, the size of the word-length effect did not change.

Recall in any serial position

For recall in any serial position, as predicted by Hypothesis 1.1, there was a phonological similarity effect, $F(1,31) = 16.52$, $p < .001$, $MSE = .25$. As predicted by Hypothesis 1.2, there was neither a main effect of task-type, $F(1,31) = 1.66$, $p = .207$, $MSE = 1.06$, nor an interaction between phonological similarity and task-type, $F(1,31) = 0.54$, $p = .470$, $MSE = .16$.

Hypothesis 1.3 predicted a word-length effect, which the data showed, $F(1,31) = 30.83$, $p < .001$, $MSE = .24$. There was a statistical trend toward more words in the simple-span task being recalled than in the complex-span task, $F(1,31) = 3.94$, $p = .056$, $MSE = 1.43$, and there was an interaction between word-length and task-type, $F(1,31) = 9.85$, $p = .004$, $MSE = .18$ (Table 36; Figure 15). Post-hoc paired t-tests showed this interaction to be due to lower recall of 3-syllable words in the complex than in the simple-span task, $t(31) = 3.15$, $p = .004$ two-tailed, but with equivalent recall of control words in both simple and complex-span tasks, $t(31) = .76$, $p = .451$ two-tailed. That is, the magnitude of the word-length effect increased in the complex-span task.

The consensus for Experiment 7 regarding Hypotheses 1.1 to 1.4 was quite clear. As predicted, a phonological similarity effect occurred in both simple and complex-span tasks. Of most relevance to Experiment 7 was that this phonological similarity effect was unaltered by concurrent sentence verification. This important result suggests independence between the processes responsible for producing the phonological similarity effect and the process allocating time to verification and rehearsal.

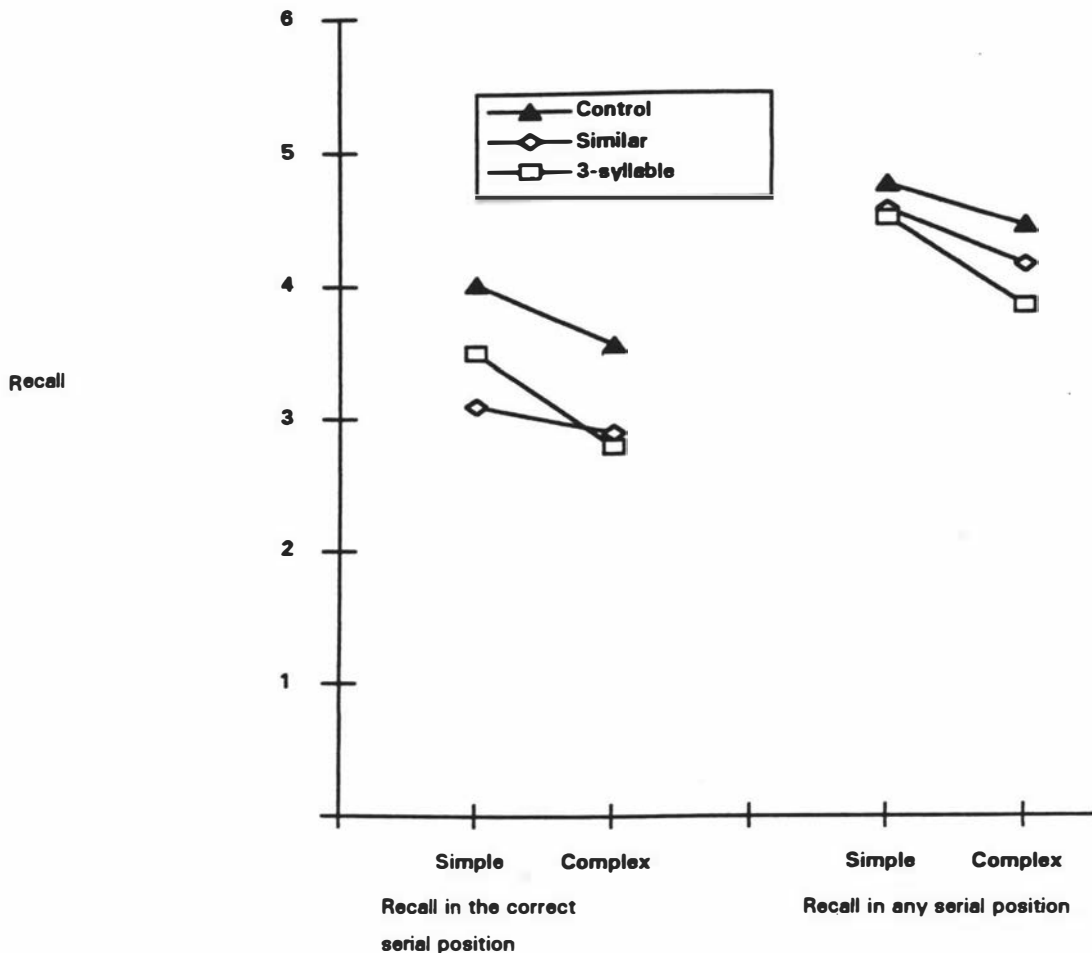


Figure 15. Recall for Experiment 7 showing the comparison between simple-span and complex-span tasks for each stimulus-type.

There was also a word-length effect established in Experiment 7. However, the data showed that this word-length effect altered with task-type. Specifically, the word-length effect increased in magnitude from simple-span to complex-span tasks for recall in any serial position (c.f., La Pointe & Engle, 1990). The present finding was consistent with Baddeley's (1986, 1992a, 1992b) model of the Articulatory Rehearsal Process which would predict that with lower overall recall (indicating more iterations of the Articulatory Rehearsal Process and increased overall decay of stimuli), the word-length effect should be larger.

Articulation rate and recall

There was one final issue to examine: did the articulation rate of the words influence recall of those words? In Experiments 5 and 6, articulation rate was not an especially potent predictor of subsequent recall.

There was a difference in articulation rate between the control (Mean = 2.08, SD = 0.41) and phonologically similar (Mean = 1.90, SD = 0.36) words $F(1,31) = 12.34$, $p = .001$, $MSE = .04$. However, regression equations failed to establish any predictive relationship between articulation rate and recall for the phonologically similar words. It could not be established that a difference in articulation rates between control and phonologically similar words has any relationship to the differences in recall between those stimulus-type conditions. Thus, Hypothesis 1.5 that articulation rate had no influence was confirmed.

In the present Experiment articulation rate for control words (Mean = 2.08, SD = 0.41) was not significantly different than for 3-syllable words (Mean = 2.03, SD = 0.36). Hence, the absence of a difference in articulation rate cannot be used to predict differences in recall between control and 3-syllable words, failing to confirm Hypothesis 1.6. In summary, articulation rate was unable to predict either the word-length or phonological similarity effects in Experiment 7. Again, as noted in Experiment 6, this failure to find articulation rate differences when a word-length effect was present raises a question as to how the word-length effect is produced.

Order measures: Order errors

Hypothesis 2.1 stated that order errors for phonologically similar words would be greater than in the control condition. Hypothesis 2.2 predicted that order errors would be equivalent for the control and 3-syllable word conditions. Data from Experiment 7 showed that there were more order

errors in the phonologically similar condition than in the control condition, $F(1,31) = 22.81$, $p < .001$, $MSE = .25$, and that no difference existed between 3-syllable and control condition order errors, $F(1,31) = 2.95$, $p = .096$, $MSE = .27$. Additionally, there were no main effects of task-type, $F(1,31) = 2.03$, $p = .164$, $MSE = .44$, nor any interactions between word-length and task-type, $F(1,31) = 1.96$, $p = .171$, $MSE = .21$. In summary, task-type had no effect on order errors; word length had no effect on order errors; but phonological similarity increased order errors (Table 37).

Table 37. Mean (SD) order errors by task-type (simple or complex) and stimulus-type (control, phonologically similar, 3-syllable; $N = 32$) for Experiment 7.

	Order errors
Simple-span	
Control	0.73 (0.57)
Similar	1.34 (0.63)
3-syllable	1.01 (0.77)
Complex-span	
Control	1.02 (0.77)
Similar	1.26 (0.70)
3-syllable	1.03 (0.59)

As was done in Experiments 5 and 6, transposition errors were analyzed in order to examine whether the nature and frequency of errors differed across the three conditions in Experiment 7. In Experiment 7, the frequency at which items were placed in the correct serial position was greater in the control condition (frequency = 1789) than in either the phonologically similar (frequency = 1650) condition, $\chi^2(1, N = 32) = 5.67$, $p = .017$, or 3-syllable (frequency = 1604) condition, $\chi^2(1, N = 32) = 10.18$, $p = .001$. Consequently, the frequency of order errors was adjusted to account for the different frequencies of correct placements by dividing the frequency of order errors into the total number of order errors and then multiplying this proportion by the frequency at which items were placed in the correct serial position. This transformation equated transposition errors for the

frequency of correct placements. Thus, any difference in transposition errors cannot be argued as resulting from differences in the frequency of recall (a content measure). With the adjusted frequencies, there was a significant difference at the $p < .001$ level, $\chi^2(1, N = 32) = 27.70$, in the overall frequency of transposition errors between the control (frequency = 489) and the phonologically similar (frequency = 665) items, but not between control and 3-syllable (frequency = 518) items. However, there was no difference in the way the transposition errors were distributed across any conditions.

Overall, Experiment 7 provided some convincing data that the *phonological similarity effect is produced by a greater loss of order information for phonologically similar than for phonologically dissimilar words*. In contrast, the loss of order information appears invariant as a function of word-length. Together, these two conclusions can be taken as validation of the utility of fractionating working memory (Baddeley, 1986, 1992a, 1992b). Finally, it would appear that the process in working memory that produces order errors is independent of the process that allocates processing to both verification and recall. In Baddeley's (1986) language, the Phonological Store is independent of the Central Executive.

Processing time : Stimulus viewing and verification times

Having examined content and order issues, attention is now turned to the time each person spends viewing the stimuli. The fundamental issue was whether more viewing time was spent on complex than simple-span tasks. Analysis of the data showed that only for 3-syllable words were viewing times longer in the complex-span than the simple-span task, $t(31) = -2.35$, $p = .025$ two-tailed.

Planned MANOVAs examining the phonological similarity effect found no overall differences between task-type, $F(1,31) = 2.33$, $p = .137$, $MSE = 2,765,611$, nor word-type, $F(1,31) =$

3.42, $p = .074$, $MSE = 65,573$, and no interaction between task-type and word-type, $F(1,31) = .54$, $p = .470$, $MSE = 59,670$.

For the word-length effect, 3-syllable words were viewed for longer than the control words, $F(1,31) = 33.88$, $p < .001$, $MSE = 110,736$. There was an effect of task-type at the trend level of significance ($p < .1$), $F(1,31) = 3.60$, $p = .067$, $MSE = 3,074,884$. Also, there was an interaction between task-type and word-length, $F(1,31) = 17.31$, $p < .001$, $MSE = 99,201$, attributable to the longer viewing times for 3-syllable words in the complex-span task but with no difference between complex and simple-span tasks in the control condition. To summarize, for the phonological similarity effect, viewing times were not different across conditions. For the word-length effect, viewing times were longer for 3-syllable than for 1-syllable control words and even longer again when the 3-syllable word was in a complex-span task (Table 38).

Table 38. Mean (SD) viewing and verification time (ms) by task-type (simple or complex) and stimulus-type (control, phonologically similar, 3-syllable; $N = 32$) for Experiment 7.

Viewing and verification time	
Simple-span	
Control	2412 (1129)
Similar	2296 (912)
3-syllable	2522 (1089)
Complex-span	
Control	2829 (1031)
Similar	2777 (1039)
3-syllable	3403 (1406)

The crux of Hypothesis 2.3 was whether viewing time differences contributed to improved recall, in the form of a speed-accuracy tradeoff. As Figure 16 shows, the natural logarithm of recall in the correct serial position was unrelated to viewing times and there were no differences in variability between task-type or between stimulus-type. Individual regressions of the natural logarithm of viewing

times against recall in the correct serial position supported this impression and provided no basis to conclude that viewing time and recall were related (Hypothesis 2.4). Thus, the data were consistent with the interpretation that, within the limited duration of working memory, more viewing time has no significant relationship to levels of recall.

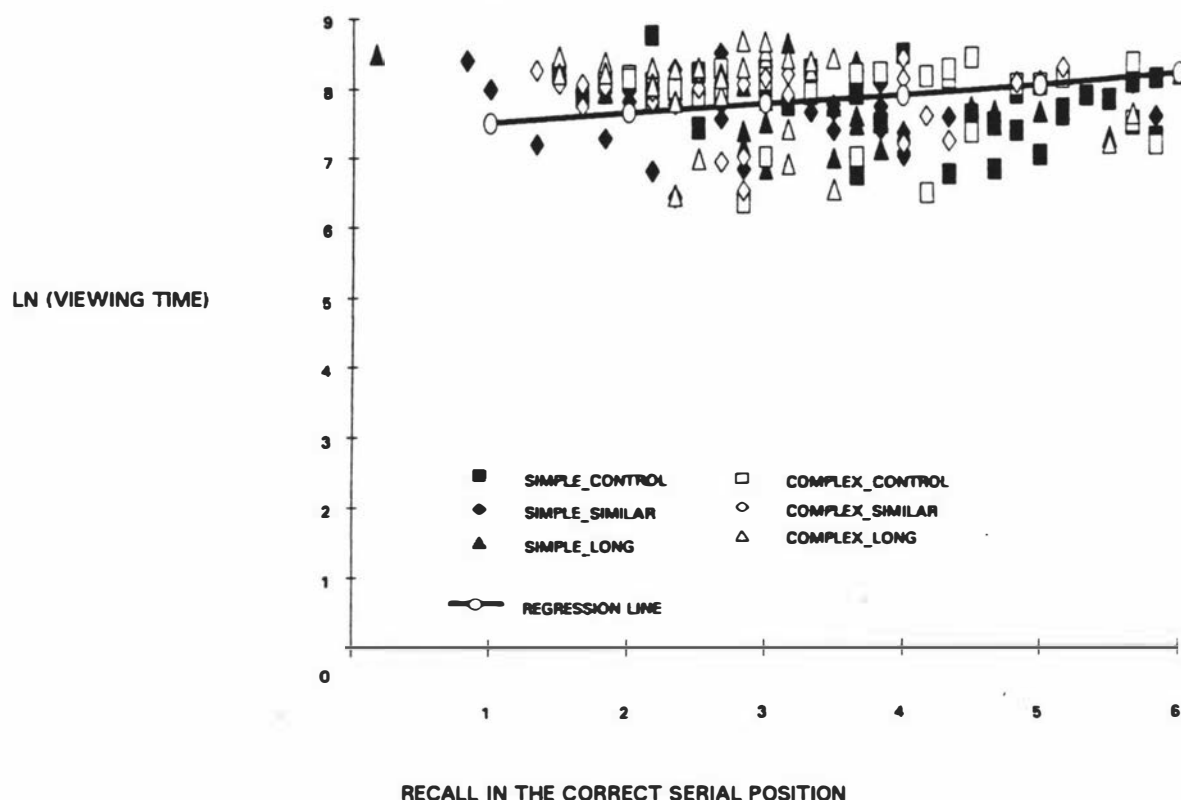


Figure 16. The natural logarithm of viewing times of each stimulus plotted against recall in the correct serial position for each task-type and for each stimulus-type.

On-line meta-memory: Pre-estimates of recall in any serial position

The accuracy of pre-estimates of recall in any serial position can be determined by (1) analyzing whether the pattern of the effects is the same for pre-estimates as for performance and (2) examining the relationship between pre-estimates and performance.

Hypothesis 3.1 predicted that persons would not be able to predict the phonological similarity effect. Contrary to this

hypothesis, persons did predict lower recall of phonologically similar than of control words, $F(1,31) = 9.83$, $p = .004$, $MSE = .10$. Persons also gave higher recall expectancies for words in the simple-span task compared to words in the complex-span task, $F(1,31) = 92.11$, $p < .001$, $MSE = .17$. In summary, persons were able to predict the pattern of the phonological similarity effect (Table 39) but incorrectly expected that they would recall more words in the simple-span than the complex-span task.

Table 39. Mean (SD) pre-estimates by task-type (simple or complex), and stimulus-type (control, phonologically similar, 3-syllable; $N = 32$) for Experiment 7.

	Pre-estimates
Simple-span	
Control	4.13 (0.97)
Similar	3.99 (0.91)
3-syllable	3.76 (0.88)
Complex-span	
Control	3.44 (0.88)
Similar	3.22 (0.84)
3-syllable	2.85 (0.68)

Hypothesis 3.2 predicted that persons would predict the word-length effect. The data showed that persons were indeed able to predict the direction of the word-length effect, $F(1,31) = 43.73$, $p < .001$, $MSE = .17$, and also accurately predicted lower recall of words presented in a complex-span than simple-span task, $F(1,31) = 107.24$, $p < .001$, $MSE = .19$. Finally, the predicted size of the word-length effect was expected to be larger in the complex-span than simple-span task, $F(1,31) = 6.35$, $p = .017$, $MSE = .07$. Post-hoc paired t -tests showed that both the control, $t(31) = 8.28$, $p < .001$ two-tailed, and 3-syllable pre-estimates, $t(31) = 9.51$, $p < .001$ two-tailed, between simple-span and complex-span recall levels were significantly different (Table 39).

Data in Experiment 7 showed that persons were capable of predicting the pattern of both the phonological similarity

and word-length effects. However, before one can conclude that they might be monitoring working memory components, there needs to be a demonstrated relationship between pre-estimates and performance. For the simple-span task, the control, $R^2 = .42$, $F(1,30) = 21.25$, $MS_{\text{residual}} = 0.66$, phonologically similar, $R^2 = .27$, $F(1,30) = 11.31$, $MS_{\text{residual}} = 0.34$, and 3-syllable, $R^2 = .33$, $F(1,30) = 14.69$, $MS_{\text{residual}} = 0.41$, conditions all showed a predictive relationship between performance and pre-estimates at $p < .05$. Similarly, for the complex-span task, the control, $R^2 = .15$, $F(1,30) = 5.21$, $MS_{\text{residual}} = 0.82$, phonologically similar, $R^2 = .15$, $F(1,30) = 5.22$, $MS_{\text{residual}} = 0.44$, and 3-syllable, $R^2 = .13$, $F(1,30) = 4.63$, $MS_{\text{residual}} = 0.71$, conditions also showed a predictive relationship between performance and pre-estimates at $p < .05$. It was also interesting to note that the magnitude of the relationship was lower for the complex than simple-span task, dropping from predicting from 27%-42% of the variance in the simple-span task to 13%-15% in the complex-span task. This drop in magnitude suggests that persons on-line meta-memory is more effective when there are not the effects of other, possibly distracting, tasks to contend with.

In summary, both the phonological similarity and word-length effects were predicted by persons in the pattern that they occurred. Furthermore, there were also predictive relationships between the pre-estimates and performance measures indicating that persons were predicting actual working memory performance and not simply some general level of performance.

Discussion

The principal expectation of Experiment 7 was that there would be *no qualitative differences between a simple and a complex-span task across the three dimensions of content, order, and processing time*. Experiment 7 shows that over the

three areas of content, order, and processing time, that whether the target words were presented alone (the simple-span task) or in the context of an additional operation such as the sentence verification task (a complex-span task) did not alter the direction or magnitude of the phonological similarity effect. Experiment 7 also shows that presenting target words in a complex-span task did not diminish, and possibly increased, the word-length effect over presentation of target words in a simple-span task. These interpretations of Experiment 7 are consistent with the results of previous studies in the current research. That is, task complexity is independent of the qualitative aspects of the qualitative/structural model of working memory. Of greater importance is that Experiment 7 demonstrates that both simple-span and complex-span tasks produce effects described as demonstrating operation of the Phonological Loop and thus support Postulate 1 (p. 85). This statement will be discussed further in the General Discussion section.

There are, however, some cautions in accepting the above conclusions about no differences between simple and complex-span tasks. These cautions relate to the level of experimental power (Keppel, 1982), and to a possible speed-accuracy tradeoff. Inspection of the data shows that in nearly every case, simple-span was higher than complex-span performance but without reaching significance. Keppel's (1982) discussion of power suggests that asking how confident one is that the alternate hypothesis, that simple-span and complex-span tasks are different, might have been accepted at these treatment levels is an appropriate question. Analysis of parametric statistics for recall in the correct serial position and recall in any serial position, for order errors, and for processing time showed that the probability of a Type 2 error was between 0.70 and 0.99 ($\alpha = 0.01$). That is, from the data in Experiment 7, the probability that the conclusion that simple-span and complex-span tasks produced similar performance levels was incorrect is very small.

It is also possible that persons who recalled more words were also the persons who spent longer viewing (and by implication rehearsing) the target items. There were two ways to overcome this criticism. First, presentation of items could have been done at a fixed pace. This solution was impractical for the sentence verification task because the increased number of verification errors would have introduced both bias and discriminability confounds into the verification task. To pace only the simple-span task would have allowed the argument made against La Pointe and Engle (1990) that the tasks varied across at least two aspects (complexity and pacing) to be also levelled at Experiment 7. The second way to overcome the speed-accuracy tradeoff criticism is to simply examine what persons did. Analysis of the results allowed the interpretation that even when, in the longer word condition, processing times were longer, these increased processing times did not display any shared variability with measures of content. Because order errors were derived from measures of content, neither could processing time have had any significant shared variability with order errors. Thus, there was no evidence for a speed-accuracy tradeoff between measures of content and order and processing time.

After consideration of both type I and II errors in inference making from the descriptive statistics and rejection of the possibility that there was a speed-accuracy trade-off, the results suggest that there are robust reasons for accepting the null hypothesis of no difference between simple and complex-span tasks. A conclusion of no difference between simple and complex-span tasks directly addresses Postulate 1 that:

1. If both quantitative/process and qualitative/structural conceptualizations of working memory are describing the same construct, then the qualitative defining factors of the qualitative/structural model (the phonological

similarity and word-length effects) ought to also be evident in a quantitatively based model.

So the conclusion that simple and complex-span task presentation modes do not differentially alter the qualitative or quantitative aspects of working memory is warranted.

The usefulness of measures of content, order, and processing time

Experiment 7 also provided data from which it was concluded that Postulate 2 was viable, namely that:

2. The factors of order and processing time of process/quantitative models will be useful in elaborating upon the qualitative/structural model of working memory, in addition to a measure of content.

For measures of content, irrespective of task-type, phonologically similar words and longer words are less well recalled than dissimilar and shorter words -- the phonological similarity and word-length effects, respectively. Experiment 7 replicates previous research both in the present dissertation (Experiment 5) and elsewhere (Conrad, 1965; Gruneberg & Melton, 1972; Healy, 1974, 1982; Watkins, et al., 1974; Wickelgren, 1965) that suggests that the phonological similarity effect is produced by the loss of order information and not necessarily content information. Finally, processing time of the stimuli has no apparent effect on the phonological similarity effect. In contrast, more processing time is used for longer than shorter words. However, despite there being more processing time for longer than shorter words, the word-length effect was not affected or produced by this longer processing time. That is, there was no speed-accuracy tradeoff between processing time and recall in any condition. Overall, the usefulness of content, order, and processing time measures has been to allow discrimination of the qualitative word-length and

phonological similarity effects in working memory from each other.

Finally, although Baddeley (1986, 1992a, 1992b) and others have claimed that articulation rate is able to predict the word-length effect, the present results do not support this assertion. It must be noted that, where a link between articulation rate and recall has been demonstrated, there have been no regression analyses to support either the magnitude of the effect nor the usefulness of any regression slopes in predicting such a relationship. In the present research, where regression analyses have been used, no link between articulation rate and recall is present. This result has some interesting implications for any model of how the word-length effect is produced by an Articulatory Rehearsal Process which can be indexed by articulation rate. At the most basic level, it may simply be that articulation rate is not the best measure of the Articulatory Rehearsal Process, something Baddeley, et al. (1975) acknowledge.

On-line meta-memory: The ability of persons to monitor their working memory

Experiment 7 shows that persons are able to predict the direction of both word-length and phonological similarity effects. Previously, in Experiments 5 and 6, persons had only demonstrated accurate prediction of word-length effects. As a further check on the certainty with which it can be concluded that persons can predict both word-length and phonological similarity effects, a power analysis was conducted (Keppel, 1982). For both the phonological similarity and word-length effects, the confidence that a type II error had not been made was greater than 0.93 and 0.99 respectively. That is, there is a very high certainty that it was the differences between word-length or phonological similarity that persons were predicting.

So, from Experiment 7, persons again were able to demonstrate on-line meta-memory for the word-length effect. This result

replicates the result in Experiment 6 and adds weight to a conclusion that persons might be able to monitor the Articulatory Rehearsal Process. In Experiment 7, persons were also able to demonstrate on-line meta-memory for the phonological similarity effect where in Experiment 5 they had not. So, from Experiment 7, the possibility is re-opened that the Phonological Store component of working memory (Baddeley, 1986, 1992a, 1992b) might also be monitored by persons.

Summary

Experiment 7 had three major aims. The first aim was to investigate whether the simple-span tasks (typically used to operationalise the qualitative/structural model) and the complex-span tasks (typically used to operationalise the quantitative/process model) measure the same construct of working memory. In psychometric terms, are the two tasks able to demonstrate some convergent validity (Anastasi, 1982)? The short answer is that the two definitions of working memory do demonstrate some convergence. The absence of a difference between simple-span and complex-span tasks in Experiment 7 suggests that a concurrent operation in working memory does not diminish either the operating capacity (e.g., Salthouse, 1990) of working memory in toto or the qualitative processes within working memory (Baddeley, 1986, 1992a, 1992b).

The second aim of the present study was to further develop the process variables of content, order, and processing time as dependent variables useful in differentiating between qualitative aspects of working memory. Again, the results show that these three aspects of working memory do differentially describe processes of working memory. Of greatest importance is that the phonological similarity effect appears to result from loss of order information; that the word-length effect is not affected differentially by loss of order information; and that both phonological similarity and word-length effects are independent of the time to input

information. These results will be discussed further in the General Discussion section in relation to working memory.

Finally, the third aim of the present study was to replicate previous findings in this set of studies that on-line meta-memory was able to monitor the Articulatory Rehearsal Process. The present results not only replicated previous findings, but also suggest that persons might be able to monitor the Phonological Store's operation as a phonological similarity effect was also predicted. These results suggest that persons may well be able to predict the specific, and not just general, operation of working memory, and thus add support for Postulate 3.

Discussion and Conclusions

CHAPTER 11: GENERAL DISCUSSION

The main aim of the present research was to investigate the relationship between two performance-based measures of working memory, and to examine whether persons are able to accurately report on their working memory. To examine whether the qualitative/structural and quantitative/process models of working memory are assessing the same construct, two sets of findings were considered necessary. First, features of the qualitative/structural model of working memory need to be demonstrated in the quantitative/process model of working memory. Second and conversely, the dimensions used to define the quantitative/process model of working memory should be shown to be useful in the qualitative/structural model of working memory. To give the punch-line away, similar effects are found in tasks which appear to be associated with differing models. Indeed, the presence of word-length and phonological similarity effects in both types of tasks seems to implicate the process that Baddeley (1986, 1992a, 1992b) termed the Phonological Loop in both models. Viewing these data strongly, they can be taken as initial evidence for the congruity of the construct of working memory across the two approaches. A more conservative interpretation of the data suggested by Parr (1996, personal communication) is that the present data reflects involvement of the Phonological Loop in a task previously presumed primarily to involve the Central Executive. The difference between these two positions depends on both how willing one is to accept the presumption that the Central Executive is what a complex span task is measuring and the degree of involvement that one ascribes to the Central Executive in a complex versus a simple span task.

Notwithstanding these qualifications, the present data, while not allowing one to conclude what the role of the Central Executive is, certainly suggest that the Phonological Loop may hold a more prominent role in complex-span tasks than previously thought.

Second, whether self-report measures yield similar results to performance measures of working memory was also investigated. There was evidence from the present set of studies that (1) persons could generally predict their working memory operation and (2) that persons were also able to predict operation of some components of the Phonological Loop. Together, these areas present data bearing on the issue of whether working memory, as used in the present body of literature, can be thought of as a unitary construct or not.

At this point, a general review will be presented of the present research in terms of its main findings. Prior to making any inferences based upon measures of a construct such as working memory, issues of reliability are examined. Also, how sensitive and what bias persons exhibited on the complex-span task will be summarized. Having established that the measures that were used to assess working memory were reliable, that persons were sensitive, and that persons were free of bias, the unity of working memory is discussed. Finally, some implications of the results for the wider body of knowledge about working memory will be discussed along with some suggested directions for future research.

PRE-CONDITIONS FOR MEASUREMENT OF WORKING MEMORY: RELIABILITY, SENSITIVITY, AND BIAS

Reliability

A major assumption made in assessing working memory is that one is assessing a cognitive process that does not vary much in its operation over short periods of time. As stated in the theory section (above), there has been little empirical basis

for making this assumption of stability over time for the Reading Span task (Daneman & Carpenter, 1980) or, by inference, for the complex-span used in the present investigation. Consequently, the first criterion for any investigation of measures of working memory must be that the score on those measures does not vary too much over time. This is a question of test-retest reliability (Anastasi, 1982). From the first two experiments in the present series, the scores on a complex-span task showed moderate to high correlation coefficients and no differences in score variability between both sessions over two weeks. The present results concur with Tirre and Peña (1992) who also found adequate test-retest reliability for a complex-span task over time. Given the low degrees of freedom in the present design, there is an empirical basis for concluding that scores on complex-span measures of working memory remain stable over short time periods.

Sensitivity

Although Daneman and Carpenter (1980, 1983) did not originally require persons to verify the sentences in the Reading Span task, other more recent applications of complex-span procedures do (e.g., Engle, et al., 1992; Salthouse, et al., 1991). In keeping with the wider body of research, the present research required persons to verify whether a sentence was correct or muddled so as to ensure that persons were indeed engaging in a concurrent cognitive operation. However, sentence verification introduces the possibility that the veridicality of verification may covary with the other independent variables. For example, persons might expend less effort in verifying sentences when they are also required to use articulatory suppression than when not. In such an instance, the reduced effort applied to sentence verification might offset the increased demands of articulatory suppression.

To avoid this problem of confounding verification sensitivity with other experimental manipulations, signal detection methods were used to examine the ability of persons to accurately detect *correct* and *muddled* sentences across experimental conditions. Using d' (MacMillan & Creelman, 1990) as an index of sensitivity, the data show two things. First, in all conditions, persons are extremely able to detect whether a sentence is to be verified as *correct* or *muddled*. Second, there is no decrement in this high sensitivity across conditions of word-type and articulatory suppression. The conclusion of this research is that persons' ability to correctly verify sentences did not confound any of the other dependent measures. It must be noted that the terms *correct* and *muddled* differ from the more usual *True* or *False* verifications. These terms were chosen so that the "C" and "M" keys on the computer keyboard could be used, keys directly below traditionally accepted 'home keys'. In using key letters that are both meaningful to the forced choice to be made and in a familiar position, it was hoped to reduce any confounds from participants having to first learn which keys to respond with. However, it is also possible that this methodological variation of using "correct/muddled" instead of the more usual "true/false" altered the results in some way. However, given that the literature describes the complex span task as robust across its many variants, it seems unlikely that this variation in responses to the forced choice task would have had any significant effect on the results.

Bias

It might also be argued that different persons may have adopted different strategies in verifying sentences in the complex-span tasks. For example, one person may tend to press the key to verify a sentence as correct when s/he knew the sentence was correct and when it was ambiguous for him or her. A second person, being less tolerant of false positives,

may have chosen to press the key indicating that a sentence was incorrect when it was ambiguous. Finally, it may happen that a left-handed person may tend to press one key more than another. These (and many other) reasons can introduce *bias* into how a person chooses to respond. Typically there are methods to reduce the impact of this bias such as counterbalancing key allocation, randomizing subjects, and having an adequate sample size. While the present research used the last two of these methods, bias was additionally dealt with by using signal detection methods and the *c* index of bias (MacMillan & Creelman, 1990) to explicitly examine what bias was present.

When bias was examined empirically, almost no bias in responding either *correct* or *muddled* was observed. Also, the bias that did exist between persons was very low in variability. The conclusion of this dissertation is that bias is not a confounding factor upon other dependent variables.

POSTULATE 1: THE COMMONALITY BETWEEN QUALITATIVE/STRUCTURAL AND QUANTITATIVE/PROCESS DEFINITIONS OF WORKING MEMORY

The primary question asked in the present research was whether the qualitative/structural and quantitative/process models of working memory are delineating the same construct in terms of how verbal information is stored. In Chapter 5 it was advanced that if both models of working memory are describing a single construct in respect of verbal information storage then two logical hypotheses can be stated: (1) aspects of the qualitative/structural model will be present in the quantitative/process model of working memory, and (2) aspects of the quantitative/process model will be present in the qualitative/structural model of working memory. From this first statement, Postulate 1 was advanced that:

1. If both quantitative/process and qualitative/structural conceptualizations of working

memory are describing the same construct, then the qualitative defining factors of the qualitative/structural model (the phonological similarity and word-length effects) ought to also be evident in a quantitatively based model.

Operationally, Postulate 1 implies that the phonological similarity, word-length, and articulatory suppression effects definitive of the qualitative/structural model of working memory will occur in the operational definition of the quantitative/process model of working memory. That is, where the phonological similarity, word-length, and articulatory suppression effects occur in a simple-span task, they will also occur in a complex-span task.

Main Results

In comparing simple and complex-span tasks, the measure of recall referred to is recall in the correct serial position. Although this measure is inherently confounded with order information, it was used because it is most like the measures typically used to define phonological similarity, word-length, and articulatory suppression effects elsewhere (e.g., Baddeley, et al., 1984; Baddeley, et al., 1975).

Experiments 1 and 2 presented persons with words plus a concurrent sentence verification task -- a complex-span task. Experiment 1 varied the phonological similarity of the words and Experiment 2 varied the length of the words.

Experiment 1 found that phonologically similar words were better recalled than phonologically dissimilar words. A phonological similarity effect is the reverse of this finding. Experiment 3 re-presented the items to another group of participants but without the concurrent verification task. Even with this change from a complex to simple-span task, the same pattern of results were obtained as in Experiment 1. The point is that the phonological similarity effect was not

absent in Experiment 1 because of the presence of the concurrent sentence verification.

Before proceeding with discussion of the overall findings, Experiments 1 and 3 raise an important issue in terms of order errors. It was argued (above) that participants may have generated words which rhymed and these led to 'correct' responses. This raises the possibility that these generated words may have been intrusions (words not in the stimulus set). In fact, inspection of the data reveal very few intrusions in either data for similar or dissimilar words. Furthermore, the lists of rhyming words are, in most cases, almost exhaustive lists of possibilities of those rhymes. Hence, it is unlikely that intrusions could account for the higher number of order errors in the phonologically similar conditions. This issue will be discussed further below.

Experiment 2, a complex-span task varying word-length, failed to find a word-length effect. Again, the experiment was repeated (Experiment 4) without the concurrent verification task, but identical in other respects. And again, removing the concurrent task did not alter the pattern of results, demonstrating that the absence of a word-length effect in Experiment 2 was probably not because of the presence of the concurrent task.

From Experiments 1 to 4, there was evidence that an absence of phonological similarity and word-length effects was not produced by the presence of a concurrent operation. This conclusion does not allow one to state what effect the concurrent operation will have when both phonological similarity and word-length effects are present. Hence, while the first four experiments provide circumstantial evidence that whether a task is complex or simple in nature may not produce qualitative changes in recall, there was clearly a need to show this null effect of complexity when both phonological similarity and word-length effects were present.

After Experiments 5 and 6 had established the parameters necessary for replication of the phonological similarity and word-length effects respectively, Experiment 7 directly compared those effects across a simple and a complex-span task. The data from Experiment 7 show that phonological similarity and word-length effects are present in both simple and complex-span tasks. Whether a task was simple or complex had no impact on the phonological similarity effect, which was present in both operationalisations. The word-length effect was also present in both simple and complex tasks, although overall there was greater recall for simple than for complex-span stimuli (Figure 3, p. 87).

Limits of the results

The first limitation of the present results is the choice of score. In using recall in the correct serial position there is a confounding of what persons recalled with what position items were recalled in. However, given that the effects under investigation have been previously demonstrated using a content plus order measure, this would not appear to invalidate the results.

Second, in using a fixed-span of six items, the present research differs from the more typical incrementing-span task used in complex-span tasks. While presenting a fixed-span task was critical to calculation of order errors, a significant focus of the present research, it must be borne in mind that presenting six items was but one span length that persons would receive in a more traditional complex-span task. However, although other span lengths were not investigated here, this does not invalidate the results for the comparison for a span of six items. Rather, it suggests that a specific comparison be conducted between incrementing and fixed-span scoring methods. In the present instance however, examination of methods used by Baddeley, et al.

(1975), suggested that differences between the two methods, if they exist, were likely to be minor.

A further limit of the present results and methodology is that the role of the Central Executive is not directly addressed. As noted earlier, it is possible that the complex span task also uses the Central Executive to a greater extent than a simple span task, a position the received wisdom has taken. However, the present data, while not precluding this possibility, suggest that the gain in recall from involvement of the Central Executive would appear to be small. This suggests that one cannot necessarily claim that a complex span task uses the Central Executive alone or indeed even primarily. The less conservative conclusion in the present dissertation is that *both* simple and complex span tasks use the Phonological Loop to varying degrees. To investigate the degree of involvement of the Central Executive and Phonological Loop, it would be necessary to systematically load the storage and processing aspects independently and to note how recall was affected. This type of manipulation would extend the certainty with which one could comment on the relative involvement in general of the Central Executive in a complex span task, a task definitive of the quantitative/process approach to working memory. If nothing else, the present research provides evidence that this type of manipulation is necessary and that it is now less prudent to simply assume that the complex span task primarily involves the Central Executive (e.g., Just & Carpenter, 1992).

The final limit of the present results relates to the effect of stimulus presentation mode on the processing of verbal information (W.V. Parr, 1996, personal communication). Thus, from a more conservative position, the present results could be argued to be limited to only self-paced and visually presented verbal items. The present work attempted to deal with the issue of pacing by examining what occurred rather

than controlling presentation time (see below). In retrospect, it is apparent that this method confounded whether stimuli were self or experimenter paced with presentation time (N. Cowan, 1996, personal communication). In terms of having visually presented verbal items, I would also remind the reader that, as outlined earlier, visual items were chosen to avoid a confounding of results from involvement of the articulatory loop which is presumed to have obligatory access to orally presented verbal items. In sum, while pacing needs further investigation and presentation mode may do, this is essentially an issue of experimental scope and does not appear to nullify the results. Rather, the reader is cautioned against assuming without good reason that these results apply irrespective of presentation mode.

The Context of The Present Results in the Literature

The finding that there was no difference in the phonological similarity effect between simple and complex-span tasks appears to be without precedent in the existing literature. Given that the phonological similarity effect is a major indicator used to assess the Phonological Store, the implication of this finding is that the Phonological Store (Baddeley, 1986, 1992a, 1992b) is almost unaffected by a concurrent cognitive operation and in fact operates equally in tasks operationally defining both the qualitative/structural and quantitative/process models of working memory.

The lower recall of complex-span than of simple-span items in the present word-length comparison is congruent with the work of La Pointe and Engle (1990). Thus, the present data suggest that the word-length effect is relatively unaffected by the concurrent operation. However, in La Pointe and Engle's first three experiments, there was a statistically *smaller* word-length effect with a complex-span than with a simple-span

task, whereas, in the present research the magnitude of the word-length effect increased slightly. Also, La Pointe and Engle's results were typically lower in magnitude (Complex, 26%-39%; Simple, 31%-57%) than the present results (Complex 47%-59%; Simple 58%-67%). One possible explanation for the differences in word-length effect magnitude is that the word-length effect is largest for recall that approaches neither a floor nor a ceiling level. As recall increases or decreases, the word-length effect becomes smaller. That is, the presence of the word-length effect is limited in a regular way by ceiling and floor effects. This explanation in no way impinges on the conclusion in the present instance that the word-length effect is separate from the effect of a concurrent operation. That is, a concurrent operation only affects the word-length effect to the extent that it raises or lowers *absolute* recall toward a ceiling or floor. Thus, the data lead to the conclusion that both simple-span and complex-span models of working memory are using the same process and in fact are addressing the same construct.

What then is the role of the concurrent operation? More specifically, how is it that a complex-span task usually predicts other cognitive processes such as reading better than a simple-span task does (Daneman & Carpenter, 1980, 1983; Dixon, Le Fevre, et al., 1988; Masson & Miller, 1983; for exceptions see La Pointe & Engle, 1990; Light & Anderson, 1985)? While this question is not a specific focus of the present research, hence precluding an empirically based answer, one hypothesis is that the concurrent operation adds explanatory power by providing a crude measure of processing speed. Salthouse and colleagues (Salthouse, 1991b, 1992, 1993a; Salthouse, & Babcock, 1991; Salthouse, & Coon, 1993) proposed that processing speed, which can be indexed by the digit-symbol subtest of the WAIS-R for example (Wechsler, 1981), contributes a significant amount of predictive power to content measures of working memory in addition to independently and directly predicting other cognitive

processes. Thus, if the concurrent operation provides an index of processing speed, then one would predict that the time spent on aspects of the operation would predict span in working memory and directly predict a cognitive task such as reading. Notwithstanding this prediction, the conclusion remains that, whatever the role of the concurrent task is, it does not appear to be as directly involved with working memory per se as had been previously assumed. Indeed, from the present results, directly measuring processing time contributed very little information as to how working memory operates.

Conclusion

The primary conclusion to be taken from the above discussion is that the presence of a concurrent task does not appear to affect the presence of phonological similarity or of word-length effects. As these effects indicate operation of the Phonological Store and Articulatory Rehearsal Process respectively (Baddeley, 1986, 1992a, 1992b), the conclusion is that both the qualitative/structural and quantitative/process models of working memory are describing the same processes (see also La Pointe & Engle, 1990, p. 1,122). This conclusion, based on empirical data, mitigates the theoretically based assertion by Just and Carpenter (1992), Gathercole and Baddeley (1993), and Hasher and Zacks (1988) that the quantitative/process model of working memory primarily involves the Central Executive, an ill-defined process (Baddeley, 1992a). In contrast, the present results suggest that both models are using the same component of working memory, the Phonological Loop, to account for verbal information storage.

In summary, the present data support Postulate 1 that qualitative/structural and quantitative/process models of working memory share common processes, with the implication that both models of working memory are addressing the same or

very similar construct. Furthermore, the data show that not only is the same construct addressed in a simple and complex-span task, but that the models are also probably using the same, not separate, parts of that construct.

POSTULATE 2: THE UTILITY OF MEASURES OF CONTENT, ORDER AND PROCESSING TIME IN DESCRIBING WORKING MEMORY

Having found that both qualitative/structural and quantitative/process models of working memory appear to use the same processes from the perspective of the qualitative/structural model, the next step was to examine whether there was an equivalent use of quantitative/process dimensions by the qualitative/structural model of working memory. The second step in the present research process then was to examine whether aspects of the quantitative/process model of working memory are relevant to the qualitative/structural model of working memory. This has been stated as Postulate 2 that:

2. The factors of order and processing time of process/quantitative models will be useful in elaborating upon the qualitative/structural model of working memory, in addition to a measure of content.

Main Results

The main results in relation to Postulate 2 will be discussed in terms of content, order, and processing time. Measures of content replicated results used to propose the qualitative/structural model of working memory. A measure of order helped further explain operation of the phonological similarity effect, but added no further information to the operation of the word-length effect. Finally, processing time was not an especially useful measure in describing working memory. However, consideration of processing time did suggest a conclusion from the data that the word-length effect was

not produced by an Articulatory Rehearsal Process, but may instead be an output process of working memory.

Content

When content was considered without regard for order or processing time, the results of the present study were as follows: phonological similarity of items reduced recall in Experiments 5 and 7 (content); increasing word-length reduced recall in Experiments 6 and 7; articulatory suppression reduced recall in Experiments 1 to 6; and articulatory suppression eliminated the phonological similarity and word-length effects in Experiments 5 and 6, respectively. Most importantly, there were no differences in the phonological similarity or word-length effects as a function of whether the task was a simple-span or complex-span one. These results have been reviewed above in the section examining Postulate 1. Thus, provided one accepts the current measures of content as valid, it would be reasonable to assert that Baddeley's (1986, 1992a, 1992b) phonological loop operates in both qualitative/structural and quantitative/process models of working memory using measures of content alone.

Thus, from the present data it is evident that those effects used to define working memory can be replicated and used in theory-making when using a content (or capacity) measure of working memory. In terms of Postulate 2, a measure of content is necessary. However, the present data, while acknowledging that a measure of content is necessary, do not allow one to accept that measuring content without considering order or processing time is sufficient to elucidate the operation of working memory. In the succeeding two sub-sections, the added contributions of order and processing time, independent of content, will be considered.

Order

The present data show that assessing order independently of content contributes toward theoretical understanding of working memory processes. That is, the present data support Postulate 2, which asserts that order is a necessary component to be used in describing the qualitative/structural model of working memory.

However, in assessing the effect of order, the data do not allow one to conclude that the Phonological Store operates as described by Baddeley (1986, 1992a, 1992b). Specifically, the present data suggest that the phonological similarity effect is produced by order confusions. That is, the more shared phonemes one item has with another, the more likely that the position of that item will not be recalled (Conrad, 1965; Gruneberg & Melton, 1972; Watkins, et al., 1974; Wickelgren, 1965). The most conservative conclusion from this result is that the phonological similarity effect occurs because of an increased loss of both item identity and item position information when items are phonologically similar.

That there was still an effect of phonological similarity with content only measures is most likely due to the content measures used in the present study being at least in part confounded with order. As persons had to recall items in the correct order primarily, it is not unreasonable to assume that the strategy used to do this influenced the measure of content. In future research, it would be important to validate this assumption through direct comparison of content and content plus order measures. In the present context, this was not done as the purpose of the research was to explore what the effect of phonological similarity was on order information.

Processing time

The final dimension used to describe working memory in the present study was processing time, defined as the time spent viewing each stimulus (where *stimulus* represents either a single word or a sentence plus word). Rather than immediately attempting to define what processing time represents, the approach taken was pragmatic in that if processing time is unrelated to working memory, then there is little need to precisely define processing time. Instead, specification of processing time parameters should be sufficient.

There were two principal results in assessing processing time. First, in Experiment 6 there was little change in recall when persons could pace their own stimulus presentation rates, compared to when the presentation pace was set by the experimenter. Second, the data from Experiments 1 to 7 provided no evidence of a relationship between the time spent processing stimuli at input and subsequent recall. Furthermore, in Experiments 6 and 7 changes in processing time were not related to the magnitude of the word-length effect. The lack of a demonstrated relationship between processing time (a time during which rehearsal is suggested as occurring in the qualitative/structural model) and word-length effects is consistent with other data (e.g., Avons, et al., 1994) in suggesting that differences in the time available for rehearsal do not correspond to changes in word-length effects.

The main limitation of the results for processing time is that it is unclear what cognitive processing was occurring during this time. Initially, it was thought (Experiments 1 and 2) that changes in processing time primarily represented rehearsal time changes. This argument was not borne out by the data which should have shown greater within-subject recall for longer rehearsal times. Instead, in one instance (Experiment 3) when articulatory suppression was introduced,

persons spent less processing time at input per item; one person described this as an attempt to see all the items before they 'disappeared'. That is, there was some sense that more time did not necessarily mean greater recall.

The implication of the present failure to find a relationship between theoretical aspects of working memory and processing time is that, provided sufficient time is available for it, working memory operates independently of whatever variations occur in stimulus input time. This result suggests that working memory is a process that occurs when sufficient time is available, but that within the parameters of 1 to 2.5 seconds, more processing time does not increase the effectiveness of working memory.

However, the conclusion that varying stimulus input times has no impact upon recall is not without empirical challenge. For example, using a moving window technique, Engle, et al. (1992), in their second experiment, found that in a sentence plus span task that viewing times for the first, last, and to-be-remembered parts of the sentence were significantly correlated with span. The middle word of the sentence did not correlate with span. Engle, et al., in their third experiment, also found that, for a simple-span task presented with a moving window technique, viewing time correlated with span for Participant-paced but not for Experimenter-paced conditions. These results do not concur with the present results in which processing time does not reliably relate to span.

There are several reasons why the conclusions by Engle, et al. (1992) were different than in the present studies. First, they used a measure of recall in any serial position with the constraint that for a trial to be counted, all items must be recalled. Thus, the scoring procedure used by Engle, et al. was quite different than in the present set of studies. Second, the moving-window technique of presentation was quite different to the presentation mode used in the present task.

Finally, Engle, et al. used 70 persons in their first two experiments and 40 in their third experiment. If the significant correlations that they obtained with this sample size were re-examined with the sample size used in the present studies, some of the correlations that Engle, et al. inferred to be significant would have failed to be significant with a smaller sample size. Thus, perhaps the present studies were deficient in sample size with the implication that the impact of processing time is probably quite small. Consequently, it would be prudent to increase statistical power in any further investigation of the relationship between processing time and recall. Also, while Engle, et al.'s results challenge the conclusion in the present instance that viewing time does not relate to recall, even their data do not show a powerful relationship between viewing time and recall, which would be expected if viewing time were an indicator of rehearsal time and rehearsal was critical to later recall. Future research would do well to question (a) how important rehearsal is to later recall, (b) whether rehearsal is time-limited, and (c) if rehearsal is time-limited, what the parameters are and when the limitation occurs.

Putting aside the question of whether viewing time relates to recall, the present data definitely provide no basis for concluding that viewing time *differences* contribute to either word-length or phonological similarity effects. Thus, the dimension of processing time does not appear to be directly important in describing working memory.

Limits of the results

In using the separate measure of content (recall in any serial position), a measure calculated separately from the content plus order measure, an attempt was made to separate content from order. However, in retrospect, it could also be argued that because persons knew that they were to try to

recall the order of items in addition to the content, that trying to simultaneously assess an order-free measure of content would be confounded with the strategy that a person chose. That is, persons were being asked to recall items in their correct position, and when that task was complete to then try to recall any additional items irrespective of their correct position. This task is different than being asked to simply recall as many items as possible. This criticism is, of course, quite valid and will not be disputed. The important point to consider is what functional effect this difference is likely to have on interpretation. Primarily, the effect would be that content-only scores might have been under-estimated. Under-estimation would have a net effect of perhaps hiding the magnitude of some effects. Thus, any conclusions based on content-only scores may suffer from a tendency toward false negatives; that is, such conclusions may be over-cautious.

Another potential limit of the results relates to what effect intrusions and omissions may have had on order errors. As discussed earlier, one could argue that intrusions are more likely to occur in rhyming words. If this were the case, there is the potential to confound what are termed order errors with intrusions. While specific measures of intrusions have not been presented here, it seems unlikely that intrusions alone could account for the effect of order errors. Visual inspection of the data reveals very few intrusions in any of the conditions, certainly many fewer than the number of order errors. A second concern might be that omissions were influencing the number of order errors. While omissions were not specifically measured, an index of the number of overall omissions is obtained by comparing recall in any serial position across conditions. This index shows that omission of items differs as a function of phonological similarity. Hence, it cannot be claimed that the phonological similarity effect results from confusion of order alone. Nonetheless, it is important in further

development of the conclusions offered here that future research measures the impact of intrusions and omissions on order errors. In particular, as is amplified in the future research section (below), distinguishing between positional order errors caused by a single omission and relative order errors caused by order confusion is an important next step (item position versus item order errors). It must also be noted that the phonologically similar stimuli used here could be considered different from the stimuli more usually used to demonstrate the phonological similarity effect (e.g., Coltheart, 1993). Setting aside that repeated stimuli produce a more robust phonological similarity effect than novel stimuli, the stimuli here do have a greater degree of rhyme than the more typical stimuli. Thus, in future research it will be important to bridge the gap between the stimuli used here and those more usually selected. However, it can also be argued that using stimuli not typically used helps dispel an argument that the phonological similarity effect is something peculiar to one set of items. Upon inspection of previous research, there does appear to be a very limited set of stimuli upon which some quite broad conclusions have been based (e.g., Baddeley, 1966a, 1968; Baddeley, et al., 1984; and Coltheart, 1993 all base their work on a single stimulus set with some minimal variation).

Finally, in considering the above conclusions about how recall appears unrelated to processing time, it must be noted that processing time was not specifically manipulated or controlled other than to compare paced versus unpaced conditions. The implication of this is that, while it can be concluded that processing time appeared unrelated to recall, one cannot definitively claim that processing time does not influence recall. Part of the difficulty in comparing these two models in terms of manipulating processing time resulted from very different methodological histories between the two. Typically, processing time is experimenter controlled for simple span tasks (e.g., Baddeley, 1986) and participant

paced for complex span tasks (e.g., Daneman & Carpenter, 1980). Hence, one is left with a choice as to whether to control processing time or not. As outlined earlier, a pragmatic approach was adopted in which the present research sought first to provide adequate sensitivity (the ability to detect any relationship between processing time and recall) for any effects (by allowing participant paced presentation) with the option later to apply experimental controls to increase specificity (the ability to separate effects on working memory from effects on other processes) of any effect of processing time. It is still felt that this open-ended approach contributes to existing research by its ability to examine whether processing time was in fact critical to obtaining results determinative of the phonological loop, which it did not appear to be.

However, it would be remiss to fail to consider the potential impact of not controlling presentation time. While regression analyses failed to show a relationship between processing time and recall, there are at least two points to consider here. First, a failure to show some relationship does not mean that a relationship does not exist. It is possible that moderating factors beyond simple time may have obscured any relationship. For example, if processing time were an index of the speed of encoding and lapse of attention, it is conceivable that the variability related to lapses of attention may have obscured the (linear) relationship between encoding time and recall. Second, if one accepts dual store models of memory, then there would be some point at which more processing time begins to add to more elaborated long term storage rather than short term but temporary maintenance. Hence, it can also be argued that longer processing times were in fact beginning to confuse the boundaries between working memory and long term memory. However, as stated before, these are issues of increasing the specificity of the research and are not able to be addressed

within the present framework which explicitly sought to increase sensitivity to effects of processing time.

Conclusion

From the present research, the data have shown that examining how both qualitative/structural and quantitative/process models of working memory account for verbal information storage across dimensions of content, order, and processing time contributes to the theoretical development of working memory. Separating the effects of content from order suggests that the phonological similarity effect is produced through confusions of order in addition to any effects due to content. The present data also suggest that the word-length effect occurs because of item content effects. However, unlike the assertion by Baddeley (1986, 1992a, 1992b) based on correlational analyses (Baddeley, et al., 1975), the present data do not show a reliable predictive relationship (using linear regression methods) between the time each item takes to articulate and later recall. Furthermore, when considering the impact of processing time, it became apparent that the word-length effect is probably produced at output, either in addition to or after the time-span traditionally thought of as devoted to rehearsal. When considered together, these conclusions demand a reconsideration of the well-accepted qualitative/structural model of working memory.

However, the principal positive finding of the present research in using measures of content, order, and processing time is that dimensions inherent to describing the quantitative/process model of working memory, are also able to extend understanding about the qualitative/structural model of working memory. That is, using these dimensions has helped link the two models of working memory in a theoretically and operationally meaningful way.

POSTULATE 3: ON-LINE META-MEMORY FOR WORKING MEMORY CONTENT

Given that the present research suggests some unity between the two models of working memory, comparing performance (objective) and self-report (subjective) measures of working memory seems all the more appropriate. In comparing the qualitative/structural and quantitative/process models of working memory, there has been a dominant assumption that performance-based assessment of working memory is the method of choice. However, there is also a body of research that uses self-report methods in the assessment of cognition. The second major focus of the present dissertation was to begin a process of comparison between performance and self-report methods of assessing working memory. To clarify this comparison, Postulate 3 was advanced which stated:

3. The components and not simply the general process of working memory are monitorable by both external and internal observation, thus linking subjective (self-report) and objective (performance) measures of working memory.

Main Results

Before discussing whether persons were able to predict changes in experimental conditions for the present research, an overview of the general trend across the seven experiments will be presented. When the mean recall in any serial position was regressed against mean pre-estimates for each condition, pre-estimates were found to generally underestimate recall by about 0.5 items out of 6 (8.3%), but otherwise were predictive of performance, $R^2 = .697$, $F(1,27) = 82.70$, $p < .001$, $MS_{\text{residual}} = .129$ as shown by the equation $\hat{y} = 1.110\hat{x} + 0.548$ (Figure 17). Also, Figure 17 shows that persons almost never over-estimated recall.

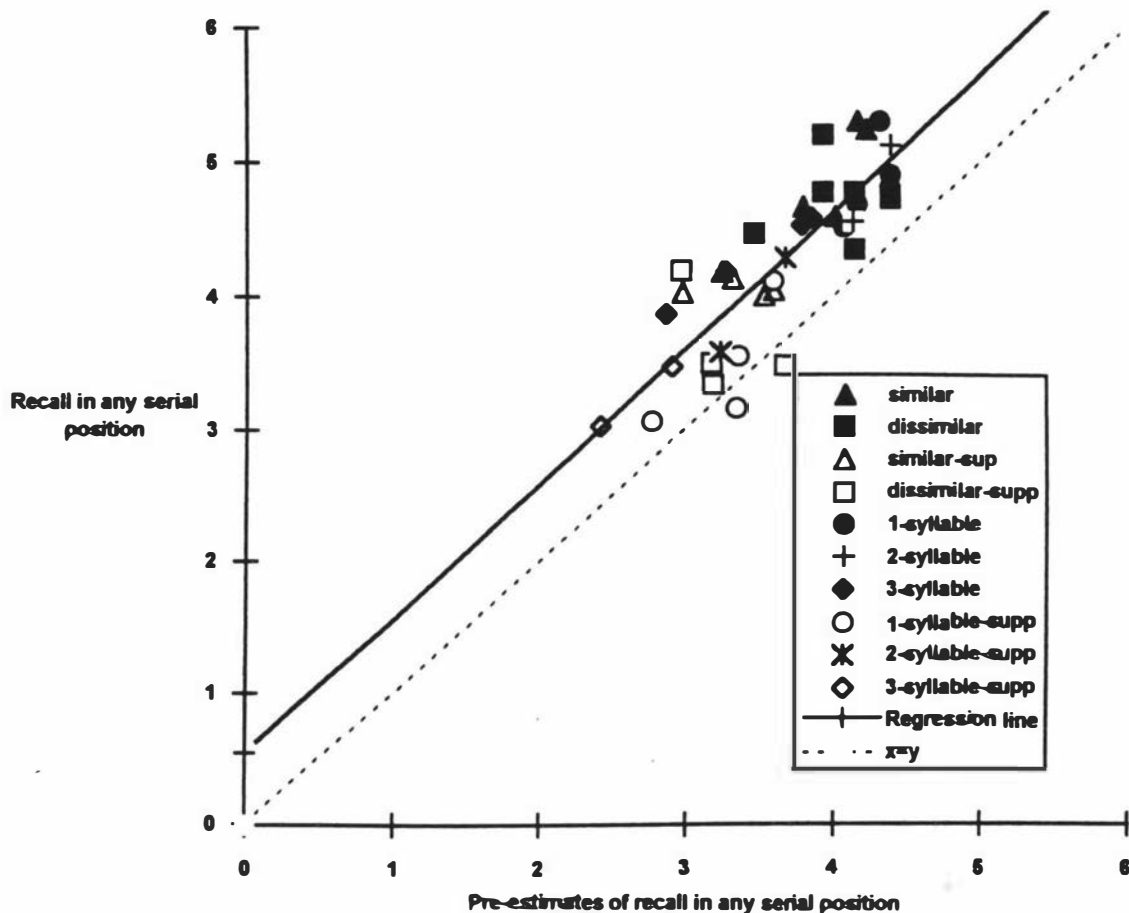


Figure 17. A regression of mean recall in any serial position against mean pre-estimates for each condition in Experiments 1 to 7.

From this general overview of the relationship between pre-estimates and recall, there is clearly an empirical basis to conclude that the general operation of working memory can be predicted by the person being assessed, although typically performance is under-estimated in magnitude, a finding replicating research on other memory tasks (Bruce, et al., 1982; Devolder, et al., 1990; Lovelace, 1984; Murphy, et al., 1981; c.f. Perlmutter, 1978). This ability to predict one's performance is in accord with Dixon's (1989) theory of meta-memory in which performance could be predicted by utilizing one's memory knowledge about (for example) how one might best recall a list of items, by one's perception of how well one did in previous trials by beliefs about immediately prior performance, and by one's overall epistemic awareness of

one's memory performance. That is, an operational on-line meta-memory measure appears to be formed by several theoretical meta-memory processes.

The general predictiveness of working memory performance by persons in the present research raises another important question: is meta-memory specific enough to predict the changes in working memory produced by variables known to change the operation of working memory? That is, are persons monitoring their working memory or are they forming an on-line estimate of working memory operation in some other way? To answer this question, persons would have to demonstrate three things in the present context: (1) that they could predict the word-length effect; (2) that they could predict the phonological similarity effect; and, (3) that they could predict the effect of articulatory suppression.

In Experiments 6 and 7, in which a word-length effect was present, persons did demonstrate an ability to differentially predict the word-length effect given only information about what parameters stimuli would vary across. Also, in Experiments 1 to 6 persons were nearly always able to predict the effect of articulatory suppression. This ability to predict both the word-length effect and the effect of articulatory suppression suggests that persons can predict the operation of the Articulatory Rehearsal Process given only information about the length of the words or whether they had to engage in articulatory suppression.

In Experiment 5, persons were not able to predict the phonological similarity effect which was observed in performance. In Experiment 7, however, persons were able to predict a phonological similarity effect through accurate predictions of recall performance. So the data were equivocal in suggesting that persons might be able to monitor the Phonological Store.

Limits of the results

The primary problem with 'internal' observation of any cognitive process is in the researcher being able to access the observations. Typically, accessing internal observations is solved either through signal detection methods (MacMillan & Creelman, 1990) or by using self-report methods (Ericsson & Simon, 1980; Gilewski & Zelinski, 1986; Herrmann, 1982; Herrmann, et al., 1986), the latter being the method used in the present research. So the first theoretical limitation on the current data is that one can never be certain how observers observe themselves or even if there is only one method of monitoring and reporting on internal processes.

Furthermore, where the present author uses the term *monitor* there is an inference that this occurs whilst working memory is operating. This raises a question of what exactly is being monitored. Are persons forming a meta-memory from the products of the working memory process or are they directly monitoring the effectiveness of actual rehearsal? Furthermore, when do the processes of working memory and of meta-working memory occur in relation to each other: Are they parallel, alternating, or contiguous processes temporally? Perhaps at the crux of these questions is the question of whether one can meaningfully divide a working memory process into storage, rehearsal, and products. Indeed, some have described memory as a vector with a type of *decoding* system for production or recall (e.g., Lewandowsky & Murdock, 1989). Until these issues are further clarified, research into meta-memory will continue to suffer from some confusion as to what exactly is being monitored. However, from the present research it would seem that some of the confusion can be reduced by (1) being clear and consistent about the theoretical basis of what one is measuring and (2) by assessing meta-memory in close temporal proximity to the beginning or end of the performance task.

Conclusion

The primary response as to whether working memory can be monitored by persons and subsequently be reported on (Postulate 3) is that it can be, at least partially. The data indicate that persons are capable of monitoring their general working memory performance and are also able to predict changes in performance in advance given only information about the type of information they will be required to recall (see also Lovelace & Marsh, 1985). If it is accepted that the word-length effect indicates the Articulatory Rehearsal Process, then the present data suggest that persons can monitor that component of working memory. However, the present dissertation suggests that the word-length effect occurs at output and may not necessarily occur because of an Articulatory Rehearsal Process as described by Baddeley (1986, 1992a, 1992b; Baddeley, et al., 1975). Hence, even though persons can predict changes in recall accurately enough to be able to predict a word-length effect, if that effect is not considered part of working memory, then one cannot reasonably claim that persons can monitor a working memory component directly.

Another reason for a cautious response to Postulate 3 is that the phonological similarity effect was predicted in Experiment 7 but not Experiment 5. So, at best the data are equivocal that persons can form accurate on-line meta-memory for recall in response to changes in stimuli across phonological similarity parameters. However, in the present dissertation, the phonological similarity effect has also been argued as resulting from order confusions. Therefore, to whatever extent persons are able to predict the phonological similarity effect, they could be said to predict the loss of order information. Although equivocal, the data do suggest that this possibility is worth further investigation. That is, in research investigating on-line meta-memory, persons were asked to predict how much content they would recall

(e.g., Bruce, et al., 1982; Devolder, et al., 1990; Lovelace, 1984; Murphy, et al., 1981; Perlmutter, 1978), but never the loss of order information. Consequently, it is unknown whether persons can monitor their effectiveness in preserving the order relations between items in their working memory.

SUMMARY

Having investigated the nature of working memory, it is now time to consider what the results of that investigation suggest about the unity of working memory, about working memory theories, and about meta-working memory.

The Unity of the Construct of Working Memory

From the present dissertation, the main conclusion is that both the qualitative/structural and quantitative/process models of working memory appear to be describing the same construct in terms of how verbal information is accounted for. Just and Carpenter (1992), Gathercole and Baddeley (1993), and Hasher and Zacks (1988) had proposed, without empirical justification, that the quantitative/process model of working memory operated in what had been termed the Central Executive. However, the present results show that both operationalisations of these models of working memory in respect of their ability to store verbal information display phonological similarity and word-length effects. Thus, one must either explain how phonological similarity and word-length effects co-occur in the Central Executive and the Phonological Loop, or infer that because those effects occur in both models that the qualitative/structural and quantitative/process models of working memory use the same processes to store verbal information. Based on the present data, the conclusion that both models utilize the same working memory processes is adopted here because there is no empirical basis to support the former conclusion and because

there is no theoretical basis to preclude the latter conclusion.

Implications of the Present Results for Theories of Working Memory

The present results also imply that one can no longer only assess capacity when measuring the operation of working memory on verbal information. The present dissertation goes beyond an approach to measuring working memory in which capacity has been perceived to be the principal dimension of interest (e.g., Cantor & Engle, 1993; Cermack, 1976; Engle, et al., 1990; Just, & Carpenter, 1992; Kershner, Henninger, & Cooke, 1984; Kyllonen & Christal, 1990; Light & Anderson, 1985; La Pointe & Engle, 1990; Turner & Engle, 1989; Whitney, et al., 1991; Wingfield, et al., 1988). In contrast, the present results show the possibility, if not benefit, of using order information in expanding understanding and description of working memory. However, as noted by Burgess and Hitch (1992), Baddeley's (1986, 1992a, 1992b) model of working memory does not account for how order information is preserved, and, from the present results, certainly is not developed enough to be able to describe how the phonological similarity effect is produced. Hence, based on the present research alone, using the phonological similarity effect to delineate the Phonological Store is at best misleading because that store would appear to be concerned with order information as much as with what is to be recalled (see also Richardson, 1984). To summarize, neither the qualitative/structural or quantitative/process models of working memory are sufficiently well developed to account for the present results when order information is considered in addition to content information.

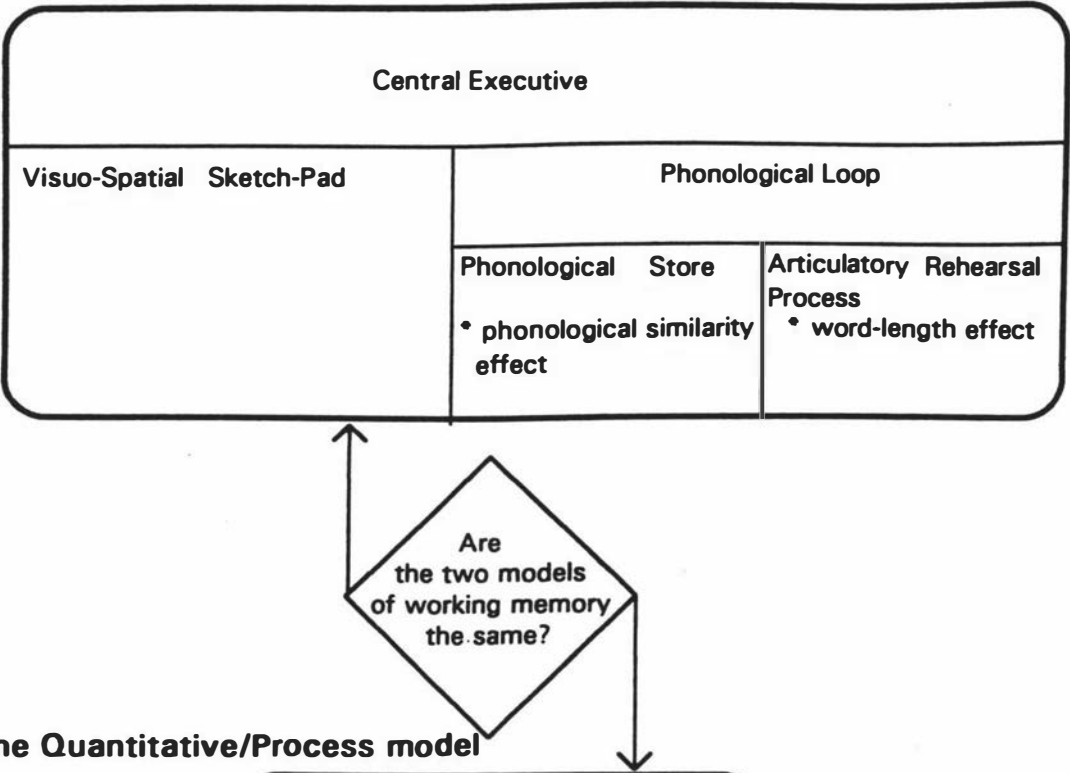
Furthermore, the present research is consistent with other research (e.g., Avons, et al., 1994; Caplan, et al., 1992; Caplan & Waters, 1994; Coltheart, et al., 1990; Cowan, 1992; Cowan, et al., 1992; Henry, 1991; Monsell, 1987; c.f.,

Baddeley & Andrade, 1994) in suggesting that the word-length effect, a second underpinning effect of the qualitative/structural model of working memory, may not be produced by the Articulatory Rehearsal Process as originally described. Specifically, the word-length effect appears unrelated to processing time at stimulus presentation, a time when one would be rehearsing items. As such, the word-length effect may not be the product of a rehearsal process. Instead, the balance of evidence from other studies favors an interpretation of the word-length effect as a product of a speech output buffer, an interpretation more consistent with the present results. This speech output buffer has been depicted as separate from the phonological input buffer, thus suggesting at least two facets to the Articulatory Rehearsal Process (Coltheart, et al., 1990; Monsell, 1987).

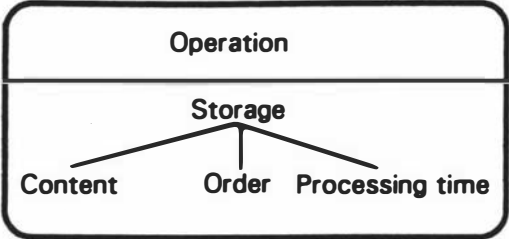
Having stated that a strong interpretation of the data in the present dissertation supports a unitary model of how working memory accounts for verbal information and that this model can be more completely described using dimensions of content and order, one must consider how well existing models of working memory accommodate these findings. In their original forms the qualitative/structural and quantitative/process models of working memory can be diagrammatically represented as in Figure 18. In its original form, the qualitative/structural model has been described as consisting of a Central Executive which monitors the operation of the Phonological Loop and the Visuo-Spatial Sketch-Pad (Baddeley, 1986, 1992a, 1992b). The Phonological Loop, the focus of the present set of studies, has been principally defined based on the phonological similarity, word-length, and articulatory suppression effects. The second model of working memory investigated in the present set of studies was the quantitative/process model. The quantitative/process model was described as consisting of operational and storage capacities varying across dimensions of content, order, and processing time (Figure 18).

Figure 18. The qualitative/structural and quantitative/process models of working memory.

The Qualitative/Structural model



The Quantitative/Process model



The present results suggest that, at the very least, the quantitative/process model involves considerably more of the phonological loop of the qualitative/structural model than previously thought. This is different than the received wisdom which states that the quantitative/process model involves the Central Executive and, by implication, not the Phonological Loop (e.g., Gathercole & Baddeley, 1993; Hasher & Zacks, 1988; Just & Carpenter, 1992). It is more likely that the quantitative/process model might be better seen as involving the Phonological Loop primarily plus some additional resources from the Central Executive to enable 'switching' between verification and recall tasks. That is,

the quantitative/process model would appear to be able to be subsumed within the qualitative structural model.

The second major challenge to both existing models from the present results is that neither of them can adequately account for order information of verbal stimuli. Thus, at a fundamental level, there is some reason to consider that there is the possibility of describing one coherent model of working memory, a model that must also account for order information and sentential input in addition to content information.

If the criteria that any unitary model of working memory in respect of verbal information must account for (a) the preservation of order information and (b) must also explain the phonological similarity and word-length effects are applied to existing working memory models, then the only model that might explain the present results is the network model of working memory proposed by Burgess and Hitch (1992). While a complete description of this model is beyond the scope of the present dissertation, the more relevant aspects of its operation will now be presented.

Burgess and Hitch (1992) began with the aim of developing a verbal working memory model, using a network modelling paradigm, that would explain serial order effects. The basic unit modelled is the phoneme. When an item is received, both input phoneme nodes and input context nodes are excited. Together, the phoneme and context nodes can be combined to represent word nodes and are stored as a set of weighted Hebbian connections. When the word node is selected, an articulatory output code is produced which, when completed, provides feedback to the item and context nodes, thus reducing the priority of those nodes for rehearsal or output. The key aspect of this network model is that context, or item order, is relevant to the eventual selection and production of a word. That is, serial order can be accounted for by this model. This mechanism is in some ways comparable to the

distinction between item identity and position in the revised model suggested below.

When the mathematical formulae it is based on are instantiated into a computer simulation, the network model of working memory shows a larger increase in order errors than in item errors as items become more phonologically similar (N. Burgess, personal communication, June 20, 1994; Burgess, 1994; Burgess & Hitch, 1992). That is, as the number of common phonemes increases, the greater the likelihood is that the item context (order) of that phoneme is substituted with another similar phoneme. Furthermore, the model is also able to predict word-length effects as a function of the number of phonemes a word contains¹, and thus further increases the convergence of the present results with their model (Burgess & Hitch). That is, when the proposed relationships between phonemes and their context are modelled, phonological similarity effects and word-length effects are produced at output.

However, one main point to be taken from the present studies is that a model of working memory must be able to account for sentential input and explain how both complex-span and simple-span tasks are able to produce equivalent effects on recall and order information. Also, a model of working memory must be able to explain visuo-spatial working memory as was done by the original qualitative/structural model (Baddeley, 1986). In this regard, the Burgess and Hitch (1992) network model of working memory appears able to explain how the Phonological Loop may operate, but is simply not as inclusive as the original qualitative/structural model.

¹In this respect, the Burgess and Hitch (1992) model implies that it is the number of phonemes, not articulation time, that determines the word-length effect. This implication that it is structural, not temporal, features has been supported by Caplan, et al. (1992) and Caplan and Waters (1994). However, Caplan and colleagues go further and suggest that it is not phonemes per se, rather it is how those phonemes are implemented as articulatory gestures that determines the word-length effect.

In conclusion, the main implication of the present results for theories of working memory are that any theory must incorporate both dimensions of content and order and must be able to describe both qualitative/structural and quantitative/process models of working memory equally well in respect of how verbal information is stored. It was shown that there currently exists only one model that begins to meet these criteria: the Burgess and Hitch (1992) network model of working memory.

Implications of the Present Results for Meta-Memory Research

The present dissertation also examined whether persons were able to directly monitor the operation of working memory. Persons were found to be able to predict a general level of working memory performance and they were partially able to differentiate between experimental conditions. When discussing meta-memory, the present results show that persons can predict changes in performance on the basis of knowledge of parameter changes. There are thus two main implications of the present research for meta-memory. First, the present results imply that one can meaningfully ask questions about persons' general and specific meta-memory for working memory. One might even refer to *meta-working memory*. The investigation of meta-working memory will be primarily limited by the exactness with which working memory is defined.

Second, if subsequent replications show that working memory can be monitored, then one must also begin to consider the possibility that working memory may be responsive to intentional changes in its operation (Reisberg, et al., 1984). In the educational arena, being able to monitor and adjust operation of working memory has an important implication for teaching children to read, for example. If working memory predicts reading performance (e.g., Daneman & Carpenter, 1980, 1983; Gathercole & Baddeley, 1990a, 1990b,

1990c, 1993; Henry, 1991; Henry & Millar, 1993), then altering working memory performance may alter reading performance (e.g., Davies, 1980).

In conclusion, the present dissertation shows that persons have a general meta-working memory ability. Also, there is some evidence that persons have a more specific ability to monitor and report on working memory components. These results are consistent with the description by Nelson and Narens (1990) of a bi-directional interaction between performance and meta-memory.

FUTURE RESEARCH

As with most research, the process of finding answers to one question inevitably raises more questions. Some of these questions have been raised previously in the context of discussion about the findings of the present research. However, the following section outlines some of the more germane questions that this present set of studies have highlighted, but perhaps not answered.

At the outset, it was stated that some researchers have presumed that complex span tasks involve the Central Executive and simple span tasks the Phonological Loop (e.g., Gathercole & Baddeley, 1993; Just & Carpenter, 1992; Hasher & Zacks, 1988). The approach this dissertation took was to examine how both these tasks stored verbal information in relation to each model of working memory. It has been argued that it would also be germane to specifically examine the role of the Central Executive and changes in how highly it is "loaded up" on recall and subsequent results (Parr, 1996, personal communication). Given that the present dissertation has identified that tasks previously presumed to occur because of Phonological Loop involvement continue to occur in a primarily Central Executive task (if one accepts that presumption), then an extension to this set of results is to examine the degree to which these Phonological Loop tasks

occur with experimentally manipulated levels of Central Executive involvement.

The present set of studies were specific in their focus on working memory for visually presented word stimuli. However, given that the original model of working memory was based on data from both auditorily and visually presented stimuli, the replication of the present results with an auditory mode of presentation might be worthwhile. This replication might take the form of the examination of the differential loss of content and order information when the phonological similarity and the length of words are manipulated, for both simple-span and complex-span tasks. However, given that Baddeley (1986) has suggested that the only difference between auditorily and visually presented stimuli would occur when articulatory suppression was used and that this difference would be that the phonological similarity effect would remain with articulatory suppression for auditorily presented words, there seemed little theoretical justification to depart from visually presented stimuli in the present instance. Related to this issue of presentation mode, it was noted earlier that a limit of the results is the confounding of pacing with presentation time (N. Cowan, 1996, personal communication). Thus, in some future research, there is reason to carefully consider whether it is presentation time that has little effect on recall or whether, when presentation time is equated, self or experimenter pacing of presentation affects recall.

A second extension to the present research would be to further develop the reasons for the phonological similarity effect. The present research shows that an increased loss of order information as a result of increasing phonological similarity does alter recall. However, what remains unclear is the extent to which the effect of order produces the phonological similarity effect independent of any effects due to item content. It might be useful to examine the extent to

which item content and item position contribute to the phonological similarity effect. Tentatively, some research using free-recall methods suggests that the phonological similarity effect might not occur or be very slight when content-only measures are used (e.g., Burgess & Hitch, 1992; Richardson, 1984) and so implies that the phonological similarity effect is primarily produced by the loss of item position information.

In addition to identifying the relative contributions of item identity and item position, future research might also make a distinction between *item position* and *item order*. In retrospect, the present research assumed that item position and item order were the same. While this assumption is not unreasonable in the light of knowledge prior to conducting the present research, it could be argued that item position relates to absolute serial position, whereas item order relates to relative position in a sequence. That is, it is possible that more order information is lost from recent than from earlier items (e.g., Burgess & Hitch, 1992), with the hint that maybe item position, not item order is critical to the phonological similarity effect. From analysis of item transpositions in the present research, the data show that the distribution of the loss of item position information is unchanged as a function of phonological similarity, that more item position information is lost for phonologically similar than phonologically dissimilar items, and that the greatest loss of item position information occurs for the middle items in a sequence. In contrast, analysis of item order (the distance from its original serial position that an item was transposed) in Experiments 5 and 6 suggests that the typical distance an item is transposed is to its immediately adjacent serial position. Thus, from these and other data (e.g., Healy, 1974), it could be hypothesized that it is item position, more than item order, that declines when order errors increase.

The present research also found some evidence that persons were able to monitor the operation of their working memory, at least in terms of item identity. However, while the present set of studies, primarily investigatory, have provided some hint that coining the term *meta-working memory* is empirically justified, the next step in examining *meta-working memory* is surely to replicate the present findings. Beyond replication, future research might begin to consider why persons under-estimate their performance. Also, given the increasing interest in the area of implicit memory, future research must consider the relationship of meta-working memory to implicit memory. A useful starting point in this endeavour might be to first delineate an empirically and theoretically based model of meta-memory in general beyond that described by Cavanaugh (1989), Dixon (1989), and Flavell and Wellman (1977).

Finally, the present research raises the possibility that meta-memory for the loss of order information should also be explored more carefully. Traditionally, the literature on meta-memory has examined persons' ability to monitor *what* they might recall. However, given that item position has been shown to be of some importance in determining recall performance, this result suggests that there also exists a functional imperative for persons to be able to monitor their ability to recall item position and possibly item order.

CONCLUSION

At the outset, the aims of the present dissertation were to determine whether (1) *the two models (qualitative/structural and quantitative/process) of working memory are examining the same construct in terms of verbal information storage* and whether (2) *persons can monitor their working memory*. Based on the present research, the short answer to both queries is Yes. First, with respect to verbal information processing, the qualitative/structural and quantitative/process models of

working memory have been shown to be examining the same or very similar construct. Second, persons have been shown to be able to at least partially monitor their working memory and thus they display *meta-working memory*.

The present research has also shown that both theory and measurement of working memory must incorporate more than the measurement of capacity alone. It has been shown that both content and order are critical to an understanding of working memory. Also, where much emphasis has been placed on the role of a concurrent operation in assessing working memory, the present research has failed to demonstrate that a concurrent operation is as critical as has been believed.

In conclusion, the present research provides an empirical analysis of some aspects of the state of working memory. This analysis has shed light on some outstanding issues and has identified more clearly the boundaries of the term *working memory*. To this end, this dissertation has also presented the author's view of what a revised qualitative/structural model of working memory might be like.

REFERENCES

- Agrawal, R., & Chhabra, H. K. (1989). Resource allocation as a function of metamemory, age, SES, and task complexity. *Psychological Studies*, 34, 87-92.
- Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Collier MacMillan.
- Atkinson, R. C., & Shiffrin, R. M. (1968). *Human memory: A proposed system and its control processes*. In K. W. Spence & J. T. Spence (Eds.) *The psychology of learning and motivation* (pp. 89-195). New York: Academic press.
- Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, 225, 82-90.
- Avons, S. E., Wright, K. L., & Pammer, K. (1994). The word-length effect in probed and serial recall. *Quarterly Journal of Experimental Psychology*, 47A, 207-231.
- Babcock, R. L., & Salthouse, T. A. (1990). Effects of increased processing demands on age differences in working memory. *Psychology and Aging*, 5, 421-428.
- Baddeley, A. D. (1966a). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, 18, 362-365.
- Baddeley, A. D. (1966b). The influence of acoustic and semantic similarity on long term memory for word sequences. *Quarterly Journal of Experimental Psychology*, 18, 302-309.
- Baddeley, A. D. (1966c). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology*, 18, 119-129.
- Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology*, 20, 249-264.
- Baddeley, A. D. (1984). The fractionation of human memory. *Psychological Medicine*, 14, 259-264.
- Baddeley, A. D. (1986). *Working memory: Oxford psychology series no. 11*. Oxford: Clarendon Press.
- Baddeley, A. D. (1990). *Human memory: Theory and practice*. Hove, UK: Lawrence Erlbaum.
- Baddeley, A. D. (1992a). Working memory. *Science*, 255, 556-559.
- Baddeley, A. D. (1992b). Is working memory working? The fifteenth Bartlett Lecture. *Quarterly Journal of Experimental Psychology*, 44A, 1-31.
- Baddeley, A.D., & Andrade, J. (1994). Reversing the word-length effect: A comment on Caplan, Rochon, and Waters. *Quarterly Journal of Experimental Psychology*, 47A, 1047-1054.

- Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer's disease: A longitudinal study. *Brain*, 119, 2521-2542.
- Baddeley, A. D., & Hitch, G. J. (1974). *Working memory*. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.
- Baddeley, A.D., & Hitch, G.J. (1977). Recency re-examined. In S Dornic (Ed.), *Attention and performance* (pp.647-667). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, 8, 485-493.
- Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology*, 36A, 233-252.
- Baddeley, A. D., & Lieberman, K. (1980). Spatial working memory. In R. A. Nickerson (Ed.), *Attention and Performance VIII* (pp. 521-539). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Baddeley, A. D., Logie, R., Bressi, S., Della Salla, S., & Spinnler, H. (1986). Dementia and working memory. *Quarterly Journal of Experimental Psychology*, 38A, 603-618.
- Baddeley, A. D., Logie, R., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, 24, 119-131.
- Baddeley, A. D., Thompson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, 14, 575-589.
- Baddeley, A. D., Vallar, G., & Wilson, B. (1987). Sentence comprehension and phonological memory: Some neuropsychological evidence. In M. Coltheart (Ed.), *Attention and performance XII*, (pp. 509-530). London: Lawrence Erlbaum associates.
- Baddeley, A. D., & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language*, 24, 490-502.
- Baillargeon, J. & Neault, S. (1989). Les modifications de la metamemoire reliees au vieillissement: Nouvelle evidence aupres d'un echantillon francophone. (Modifications of metamemory linked to aging: New evidence from a French-speaking. *Canadian Journal on Aging*, 8, 343-354.
- Barclay, C. R. (1981). On the relation between memory and metamemory. *Psychological Record*, 31, 153-156.
- Beatty, W. W., & Monson, N. (1991). Metamemory in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, 13, 309-327.
- Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *Quarterly Journal of Experimental Psychology*, 39A, 467-478.

- Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology*, 36, 701-711.
- Bishop, D. V. M., & Robson, J. (1989). Unimpaired short-term memory in congenitally speechless individuals: Implications for the notion of 'articulatory coding'. *Quarterly Journal of Experimental Psychology*, 41A, 123-141.
- Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, 6, 173-189.
- Brainerd, C. J., & Kingma, J. (1985). On the independence of short-term memory and working memory in cognitive development. *Cognitive Psychology*, 17, 210-247.
- Brooks, J. O., & Watkins, M. J. (1990). Further evidence of the intricacy of memory span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 1134-1141.
- Brown, A. L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10). New York: Academic Press.
- Brown, J. A. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 10, 12-21.
- Bruce, P. R., Coyne, A. C., & Botwinick, J. (1982). Adult age differences in metamemory. *Journal of Gerontology*, 37, 354-357.
- Bunnell, J. (1986). *Position and identity in visual information processing*. Unpublished dissertation. Victoria University, Wellington, New Zealand.
- Burgess, N. (1994). *A solvable connectionist model of immediate recall of ordered lists*. Unpublished manuscript, University College London.
- Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, 31, 429-460.
- Campbell, J. I., & Charness, N. (1990). Age-related declines in working memory skills: Evidence from a complex calculation task. Special Section: Cognitive training in later adulthood. *Developmental Psychology*, 26, 879-888.
- Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long term memory activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1101-1114.
- Cantor, J., Engle, R. W., & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence*, 15, 229-246.
- Caplan, D., Rochon, E., & Waters, G.S. (1992). Articulatory and phonological determinants of word-length effects in span tasks. *The Quarterly Journal of Experimental Psychology*, 45A, 177-192.

- Caplan, D. & Waters, G.S. (1994). Articulatory length and phonological similarity in span tasks: A reply to Baddeley and Andrade. *The Quarterly Journal of Experimental Psychology*, 47A, 1055-1062.
- Cavanaugh, J. C. (1989). The importance of awareness in memory aging. In L. W. Poon, B. A. Williams, & D. C. Rubin (Eds.), *Everyday cognition in adulthood and late life* (pp. 416-436). Cambridge, U. K: Cambridge University Press.
- Cavanaugh, J. C., & Borkowski, J. G. (1979). The metamemory-memory 'connection': Effects of strategy training and maintenance. *Journal of General Psychology*, 101, 161-174.
- Cavanaugh, J. C., & Murphy, N. Z. (1986). Personality and metamemory correlates of memory performance in younger and older adults. *Annual meeting of the Gerontological Society of America: The role of personality and social factors in intellectual aging* (185, New Orleans, Louisiana). *Educational Gerontology*, 12, 385-394.
- Cavanaugh, J. C., & Perlmuter, M. (1982). Metamemory: A critical examination. *Child Development*, 53, 11-28.
- Cavanaugh, J. C., & Poon, L. W. (1989). Metamemorial predictors of performance in young and older adults. *Psychology and Aging*, 4, 365-368.
- Cermack, L. S. (1976). The encoding capacity of a patient with amnesia due to encephalitis. *Neuropsychologia*, 14, 311-326.
- Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning and Verbal Behaviour*, 19, 722-735.
- Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behaviour*, 15, 17-32.
- Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term-memory recall of repeated and novel word lists. *Memory & Cognition*, 21, 539-545.
- Coltheart, V., Avons, S. E., & Trollope, J. (1990). Articulatory suppression and phonological codes in reading for meaning. *Quarterly Journal of Experimental Psychology*, 42A, 375-399.
- Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, 55, 75-84.
- Conrad, R. (1965). Order error in the immediate recall of sequences. *Journal of Verbal Learning and Verbal Behaviour*, 4, 161-169.
- Conrad, R. (1967). Interference or decay over short retention intervals? *Journal of Verbal Learning and Verbal Behaviour*, 6, 49-54.
- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology*, 55, 429-432.

- Corballis, M. C. (1966). Rehearsal and decay in immediate recall of visually and aurally presented items. *Canadian Journal of Psychology*, 20, 43-51.
- Cornolodi, C., Gobba, C., & Mazzoni, G. (1991). On metamemory-memory relationship: Strategy availability and training. *International Journal of Behavioural Development*, 14, 101-121.
- Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 142-150.
- Cowan, N. (1992). Verbal memory and the timing of spoken recall. *Journal of Memory and Language*, 31, 668-684.
- Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language*, 31, 1-17.
- Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behaviour*, 9, 143-148.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 268-294.
- Crowder, R. G. (1967). Prefix effects in immediate memory. *Canadian Journal of Psychology*, 21, 450-461.
- Crowder, R. G. (1993). Short-term memory: Where do we stand ? *Memory & Cognition*, 21, 142-145.
- Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics*, 5, 365-373.
- Czerniawska, E. (1983). Metamemory: A literature review. *Psychologia-Wychowawcza*, 26, 143-158.
- Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, 20, 445-464.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, 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.
- Davies, G. M. (1980). Can memory be educated? *Educational Studies*, 6, 155-161.

- Devolder, P. A., Brigham, M. C., & Pressley, M. (1990). Memory performance awareness in younger and older adults. *Psychology and Aging, 5*, 291-303.
- Devolder, P. A., & Pressley, M. (1989). Metamemory across the adult lifespan. Special Issue: Psychology of aging and gerontology. *Canadian Psychology, 30*, 578-587.
- Dixon, R. A. (1989). Questionnaire research on metamemory and aging: Issues of Structure and function. In L. W. Poon, B. A. Williams, & D. C. Rubin (Eds.), *Everyday cognition in adulthood and late life* (pp. 416-436). Cambridge, U. K. :Cambridge University Press.
- Dixon, R. A., Hertzog, C., & Hultsch, D. F. (1986). The multiple relationships among Metamemory in Adulthood (MIA) scales and cognitive abilities in adulthood. *Human Learning Journal of Practical Research and Applications, 5*, 165-177.
- Dixon, R. A., & Hultsch, D. F. (1983a). Structure and development of metamemory in adulthood. *Journal of Gerontology, 38*, 682-688.
- Dixon, R. A., & Hultsch, D. F. (1983b). Metamemory and memory for text relationships in adulthood: A cross-validation study. *Journal of Gerontology, 38*, 689-694.
- Dixon, R. A., Hultsch, D. F., & Hertzog, C. (1988). The Metamemory in Adulthood (MIA) Questionnaire. *Psychopharmacology Bulletin, 24*, 671-688.
- Dixon, P., LeFevre, J. A., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology, 80*, 465-472.
- Dobbs, A. R., Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging, 4*, 500-503.
- Ebbinghaus, H. (1964/1885). *Memory: A contribution to experimental psychology*. New York: Dover Books.
- Elley, W. B. (1975). *Assessing the difficulty of reading materials: The word frequency method*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Ellis, A. W. (1980). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behaviour, 19*, 624-634.
- Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology, 71*, 43-51.
- Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 972-992.
- Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. *Journal of Educational Research, 84*, 253-262.

- 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.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp 161-190). Washington, DC: Winston.
- Estes, W. K. (1973). Phonemic coding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning and Verbal Behaviour*, 12, 360-372.
- Farah, M. J., Hammond, K. M., Levine, D. N., & Calvanio, R. (1988). Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive Psychology*, 20, 439-462.
- Flavell, J. H. (1971). First discussant's comments. What is memory development the development of? *Human Development*, 14, 272-278.
- Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail, Jr., & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-34). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Foos, P. W. (1988). Adult age differences in working memory. *Psychology and Aging*, 4, 269-275. .
- Gathercole, S. E. (1994). Neuropsychology and working memory: A review. *Neuropsychology*, 8, 494-505.
- Gathercole, S. E., & Baddeley, A. D. (1990a). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, 28, 200-213.
- Gathercole, S. E., & Baddeley, A. D. (1990b). Phonological memory deficits in language-disordered children: Is there a causal connection? *Journal of Memory and Language*, 29, 336-360.
- Gathercole, S. E., & Baddeley, A. D. (1990c). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, 81, 439-454.
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working Memory and language*. Hove: U. K., Lawrence Erlbaum Associates.
- Gick, M. L., Craik, F. I. M., & Morris, R. G. (1988). Task complexity and age differences in working memory. *Memory & Cognition*, 16, 353-361.
- Gilewski, M. J., & Zelinski, E. M. (1986). Questionnaire assessment of memory complaints In L. W. Poon (Ed.), *Handbook for clinical memory assessment of older adults*. (pp. 93-107). Washington, DC: American Psychological Association.

- Gilweski, M. J., Zelinski, E. M., & Thompson, L. W. (1983). *Abbreviating the Metamemory Questionnaire: Factor structure and norms for adults*. Paper presented at the meeting of the American Psychological Association, Anaheim, CA.
- Greene, R. L. (1992). *Human memory: Paradigms and paradoxes*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424-429.
- Gruneberg, M. M., & Melton, K. C. (1972). Acoustic confusions and order forgetting. *Acta Psychologica*, 36, 48-53.
- Hanley, J. R., Young, A. W., & Pearson, N. A. (1991). Impairment of the visuospatial sketchpad. *Quarterly Journal of Experimental Psychology*, 43A, 101-126.
- Hartley, J. T. (1986). Reader and text variables as determinants of discourse memory in adulthood. *Psychology and Aging*, 1, 150-158.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.) *The psychology of learning and motivation* (Vol. 21; pp. 193-225). New York: Academic Press.
- Hasselhorn, M., & Hager, W. (1989). Prediction accuracy and memory performance: Correlational and experimental tests of a metamemory hypothesis. *Psychological Research*, 51, 147-152.
- Healy, A. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, 13, 644-655.
- Healy, A. (1982). Short-term memory for order information. In G. H. Bower (Ed.) *The psychology of learning and motivation* (Vol. 16; pp. 191-238). New York: Academic Press.
- Henry, L. A. (1991). The effects of word length and phonemic similarity in young children's short-term memory. *Quarterly Journal of Experimental Psychology*, 43A, 35-52.
- Henry, L. A., & Millar, S. (1993). Why does memory span improve with age? A review of two current hypotheses. *European Journal of Psychology*, 5, 241-287.
- Herrmann, D. J. (1982). Know thy memory: The use of questionnaires to assess and study memory. *Psychological Bulletin*, 92, 434-452.
- Herrmann, D. J., Grubs, L., Sigmundi, R. A., & Grueneich, R. (1986). Awareness of memory ability before and after relevant memory experience. *Human Learning Journal of Practical Research and Applications*, 5, 91-107.
- Hertzog, C., Dixon, R. A., & Hultsch, D. F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging*, 5, 215-227.

- Hertzog, C., Dixon, R. A., Schulenberg, J. E., & Hultsch, D. F. (1987). On the differentiation of memory beliefs from memory knowledge: The factor structure of the Metamemory in Adulthood Scale. *Experimental Aging Research, 13*, 101-107.
- Hoosain, R., & Salili, F. (1988). Language differences, working memory, and mathematical ability. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues, Vol. 2: Clinical and educational implications* (pp. 512-517). Chichester: Wiley.
- Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology, 47*, 72-87.
- Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1987). Age differences in metamemory: Resolving the inconsistencies. *Canadian Journal of Psychology, 4*, 193-208.
- Hultsch, D. F., Hertzog, C., Dixon, R. A., & Davidson, H. A. (1988). Memory self-knowledge and self-efficacy in the aged. In M. L. Howe & C. J. Brainerd (Eds.), *Cognitive development in adulthood: Progress in cognitive development research*. New York: Springer-Verlag.
- Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Memory and metamemory: Comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology, 17*, 3-11.
- Jay, G. M., & Willis, S. L. (1992). Influence of direct computer experience on older adults' attitudes toward computers. *Journal of Gerontology, 47*, P250-P257.
- Jones, D. M. (1992). Objects, streams, and threads of auditory attention. In A. D. Baddeley, & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control (A tribute to Donald Broadbent)* (pp. 87-104). London: Oxford University Press.
- Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 369-381.
- Jones, D. M., & Macken, W. J. (in press). Phonological similarity in the irrelevant speech effect: Within or between-stream similarity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*
- Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *Quarterly Journal of Experimental Psychology, 44A*, 645-669.
- Jones, D. M., Miles, C., & Page, J. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal, or memory? *Applied Cognitive Psychology, 4*, 89-108.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

- Kausler, D. H. (1991). *Experimental psychology, cognition, and human aging* (2nd ed.). New York: Springer-Verlag.
- Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Kershner, J. R., Henninger, P., & Cooke, W. (1984). Writing induces a right hemisphere linguistic advantage in dysphonetic dyslexic children: Implications for attention and capacity models of laterality. *Brain and Language*, 21, 105-122.
- Kintsch, W., & Buschke, H. (1969). Homophones and synonyms in short-term memory. *Journal of Experimental Psychology*, 80, 403-407.
- Klapp, S. T., Marshburn, E. A., & Lester, P. T. (1983). Short-term memory does not involve the 'working memory' of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General*, 112, 240-264.
- Klatzky, R. L. (1980). *Human memory: Structures and processes*. San Francisco, CA: W. H. Freeman and Company.
- Kontos, S., Swanson, H. L., & Frazer, C. T. (1984). Memory-metamemory connection in intellectually gifted and normal children. *Psychological Reports*, 54, 930.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.
- Kurtz, B. E., & Weinert, F. E. (1989). Metamemory, memory performance, and causal attributions in gifted and average children. *Journal of Experimental Child Psychology*, 48, 45-61.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity? *Intelligence*, 14, 389-433.
- La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 1118-1133.
- Lachman, M. E., & Jelalian, E. (1984). Self-efficacy attributions for intellectual performance in young and elderly adults. *Journal of Gerontology*, 39, 577-582.
- Lachman, J. L., Lachman, R., & Thronesbury, C. (1979). Metamemory through the adult lifespan. *Developmental Psychology*, 15, 543-550.
- Laughery, K. R. (1969). Computer simulation of short-term memory: A component decay model. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation*. Orlando, FL: Academic Press.
- Leonesio, R. J., & Nelson, T. D. (1990). Do different metamemory judgements tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 464-470.

- Levy, B. A. (1971). The role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, 10, 123-132.
- Lewandowsky, S., & Murdock Jr, B. B. (1989). Memory for serial order. *Psychological Review*, 96, 25-57.
- Light, L. L., & Anderson, P. A. (1985). Working memory capacity, age, and memory for discourse. *Journal of Gerontology*, 40, 737-747.
- Loewen, E. R., Shaw, R. J., & Craik, F. I. M. (1990). Age differences in components of metamemory. *Experimental Aging Research*, 16, 43-48.
- Logie, R. H. (1986). Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology*, 38A, 229-247.
- Longoni, A. M., Richardson, J. T. E., & Aiello, A. (1993). Articulatory rehearsal and phonological storage in working memory. *Memory & Cognition*, 21, 11-22.
- Lovelace, E. (1984). Metamemory: Monitoring future recallability during study. *Bulletin of the Psychonomic Society*, 22, 497-500.
- Lovelace, E. A., & Marsh, G. R. (1985). Prediction and evaluation of memory performance by young and old adults. *Journal of Gerontology*, 40, 192-197.
- MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, 24, 56-98.
- MacMillan, N. A., & Creelman, C. D. (1991). *Signal detection theory: A users guide*. New York: Cambridge University Press.
- 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.
- McGlyn, S. M., & Kasniak, A. W. (1991). When metacognition fails: Impaired awareness of deficit in Alzheimer's disease. *Journal of Cognitive Neuroscience*, 3, 183-189.
- McNamara, B. E. (1980). Implications of research on metamemory. *Academic Therapy*, 16, 133-137.
- Miles, C., Jones, D. M., & Madden, C. A. (1991). Locus of the irrelevant speech effect in short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 578-584.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Monsell, S. (1987). On the relation between lexical input and output pathways for speech. In A. Allport, D. G. Mackay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Relationships between listening, speaking, reading, and writing* (pp. 272-311). London: Academic Press.

- Morris, N. (1989). Spatial monitoring in visual working memory. *British Journal of Psychology*, 80, 333-349.
- Morris, N., & Jones, D. M. (1990a) Habituation to irrelevant speech: Effects on a visual short-term memory task. *Perception and Psychophysics*, 47, 291-297.
- Morris, N., & Jones, D. M. (1990b) Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, 81, 111-121.
- Morris, N., Jones, D. M., & Quayle, A. (1989). Memory disruption by background speech and singing. In E. D. McGaw (Ed.), *Contemporary ergonomics* (pp. 494-499). London: Taylor and Francis.
- Morris, R. G., & Baddeley, A. D. (1988). Primary and working memory functioning in Alzheimer-type dementia. *Journal of Clinical and Experimental Neuropsychology*, 10, 279-296.
- Morris, R. G., Craik, F. I. M., & Gick, M. L. (1990). Age differences in working memory tasks: The role of secondary memory and the central executive system. *The Quarterly Journal of Experimental Psychology*, 42A, 67-86.
- Morris, R. G., Gick, M. L., & Craik, F. I. M. (1988). Processing resources and age differences in working memory. *Memory & Cognition*, 16, 362-366.
- Murdock, B. B. Jr. (1962). The serial position effect in free-recall. *Journal of Experimental Psychology*, 64, 482-488.
- Murdock, B. B. Jr. (1965). Effects of a subsidiary task on short-term memory. *British Journal of Psychology*, 56, 413-419.
- Murphy, M. D., Sanders, R. E., Gabrieshki, A. S., & Schmitt, F. A. (1981). Metamemory in the aged. *Journal of Gerontology*, 36, 185-193.
- Murphy, M. D., Schmitt, F. A., Caruso, M. J., & Sanders, R. E. (1987). Metamemory in older adults: The role of monitoring in serial recall. *Psychology and Aging*, 2, 331-339.
- Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology*, 21, 263-276.
- Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, 78, 679-684.
- Murray, D.J. (1980). Research on human memory in the nineteenth century. In J.G. Seamon (Ed.), *Human Memory: Contemporary readings* (pp. 5-23). New York: Academic Press.
- Murray, D. J., Rowan, A. J., & Smith, K. H. (1988). The effect of articulatory suppression on short term recognition. *Canadian Journal of Psychology*, 42, 424-436.
- Naveh-Benjamin, M., & Ayres, T. J. (1986). Digit span, reading rate, and linguistic relativity. *Quarterly Journal of Experimental Psychology*, 38, 739-751.

- Nelson, T. O., Gerler, D., & Narens, L. (1984). Accuracy of feeling-of-knowing judgements for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General*, 113, 282-300.
- Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.) *The psychology of learning and motivation* (Vol. 26; pp. 125-173). New York: Academic Press.
- Niederehe, G., & Yoder, C. (1989). Metamemory perceptions in depressions of young and older adults. *Journal of Nervous and Mental Disease*, 177, 4-14.
- Parkin, A. J., Bell, W. P., & Leng, N. R. (1988). A study of metamemory in amnesic and normal adults. *Cortex*, 24, 143-148.
- Pastore, R. E., & Scheirer, C. J. (1974). Signal detection theory: Considerations for general application. *Psychological Bulletin*, 81, 945-958.
- Perfect, T. J., & Stollery, B. (1993). Memory and metamemory performance in older adults: One deficit or two? *Quarterly Journal of Experimental Psychology*, 46A, 119-136.
- Perlmutter, M. (1978). What is memory the aging of? *Developmental Psychology*, 14, 330-345.
- Peterson, L. R., & Johnson, S. T. (1971). Some effects of minimizing articulation on short-term retention. *Journal of Verbal Learning and Verbal Behaviour*, 10, 346-354.
- Peterson, L. R., & Peterson, M. J. (1959). Short term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193-198.
- Pressley, M., Borkowski, J. G., & O'Sullivan, J. T. (1984). Memory strategy instruction is made of this: Metamemory and durable strategy use. *Educational Psychologist*, 19, 94-107.
- Prevey, M. L., Delaney, R. C., & Matson, R. H., (1988). Metamemory in temporal lobe epilepsy: Self-monitoring of memory functions. *Brain and Cognition*, 7, 298-311.
- Rebok, G. W., & Balcerak, L. J. (1989). Memory self-efficacy and performance differences in young and older adults: The effect of mnemonic training. *Developmental Psychology*, 25, 714-721.
- Reisberg, D., Rappaport, I., & O'Shannessy, M. (1984). Limits of working memory: The digit digit-span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 203-221.
- Richardson, J. T. (1984). Developing the theory of working memory. *Memory and Cognition*, 12, 71-83.
- Richardson, J. T. E., Greaves, D. E., & Smith, M. M. C. (1980). Does articulatory suppression eliminate the phonemic similarity effect in short-term recall? *Bulletin of the Psychonomic Society*, 16, 417-420.

- Salame, P., & Baddeley A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behaviour*, 21, 150-164.
- Salame, P., & Baddeley A. D. (1986a). Phonological factors in STM: Similarity and the unattended speech effects. *Bulletin of the Psychonomic Society*, 24, 263-265.
- Salame, P., & Baddeley A. D. (1986b). The unattended speech effect: Perception or memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 525-529.
- Salame, P., & Baddeley A. D. (1987). Noise, unattended speech, and short-term memory. *Ergonomics*, 30, 1185-1193.
- Salame, P., & Baddeley A. D. (1989). Effects of background music on phonological short-term memory. *Quarterly Journal of Experimental Psychology*, 41A, 107-122.
- Salame, P., & Baddeley A. D. (1990). The effects of irrelevant speech on immediate free recall. *Bulletin of the Psychonomic Society*, 28, 540-542.
- Salisbury, D. F. (1990). Cognitive psychology and its implications for designing drill and practice programs for computers. *Journal of Computer Based Instruction*, 17, 23-30.
- Salthouse, T. A. (1985). *A theory of cognitive aging*. Amsterdam: North Holland.
- Salthouse, T. A. (1988a). Initiating the formalization of theories of cognitive aging. *Psychology and Aging*, 3, 3-16.
- Salthouse, T. A. (1988b). Resource-reduction interpretations of cognitive aging. *Developmental Review*, 8, 238-272.
- Salthouse, T. A. (1988c). The role of processing resources in cognitive aging. In M. L. Howe & C. J. Brainerd (Eds.), *Cognitive development in adulthood: Progress in cognitive development research* (pp. 185-240). New York: Springer-Verlag.
- Salthouse, T. A. (1990). Working memory as a processing resource in cognitive aging. Special Issue: Limited resource models of cognitive development. *Developmental Review*, 10, 101-124.
- Salthouse, T. A. (1991a). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum Associates.
- Salthouse, T. A. (1991b). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, 2, 179-183.
- Salthouse, T. A. (1992). What do adult age differences in the digit symbol substitution test reflect? *Journal of Gerontology*, 47, P121-P128.
- Salthouse, T. A. (1993a). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journals of Gerontology*, 48, P29-P36.

- Salthouse, T. A. (1993b). Attentional blocks are not responsible for age-related slowing. *Journals of Gerontology*, 48, P263-P270.
- Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, 27, 763-776.
- Salthouse, T. A., Babcock, R. L., & Shaw, R. J. . (1991). Effects of adult age differences on structural and operational capacities in working memory. *Psychology and Aging*, 6, 118-127.
- Salthouse, T.A. & Coon, V.E. (1993). Influence of task-specific processing speed on age differences in memory. *Journals of Gerontology*, 48, P245-P255.
- Salthouse, T. A., & Mitchell, D. R. D. (1989). Structural and operational capacities in integrative spatial ability. *Psychology and Aging*, 4, 18-25.
- Salthouse, T. A., Mitchell, D. R. D., Skovronek, E., & Babcock, R. L. (1989). Effects of adult age and working memory on reasoning abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 507-516.
- Salthouse, T. A., & Skovronek, E. (1992). Within-context assessment of age differences in working memory. *Journals of Gerontology*, 47, P110-P120.
- Schlechter, T. M., Hermann, D. J., & Toglia, M. P. (1990). An investigation of people's metamemories for naturally occurring events. *Applied Cognitive Psychology*, 4, 213-217.
- Schneider, W. (1993). Varieties of working memory as seen in biology and in connectionsit/control architectures. *Memory & Cognition*, 21, 184-192.
- Schneider, W. & Detweiler, M. (1988). A connectionist/control architecture for working memory. In G. H. Bower (Ed.) *The psychology of learning and motivation* (Vol. 21; pp. 53-119). New York: Academic Press.
- Schweikert, R., Guentert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science*, 1, 74-77.
- Shiffrin, R. J., & Cook, J. R. (1978). Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behaviour*, 17, 189-218.
- Spinnler, H., Della Sala, S., Bandera, R., & Baddeley, A. D. (1988). Dementia, aging, and the structure of human memory. *Cognitive Neuropsychology*, 5, 193-211.
- Swanson, H. L. (1993). Individual differences in working memory: A model testing and subgroup analysis of learning-disabled and skilled readers. *Intelligence*, 17, 285-332.
- Swanson, H. L. (1994). Short-term memory and working memory: Do both contribute to our understanding of academic achievement in children and adults with learning disability? *Journal of Learning Disabilities*, 27, 34-50.

- Swanson, H. L., Cochran, K. F., & Ewers, C. A. (1990). Can learning disabilities be determined from working memory performance? *Journal of Learning Disabilities*, 23, 59-67.
- Tehan, G., & Humphreys, M. S. (1988). Articulatory loop explanations of memory span and pronunciation rate correspondences: A cautionary note. *Bulletin of the Psychonomic Society*, 26, 293-296.
- Tell, P. M. (1972). The role of certain acoustic and semantic factors at short and long term retention intervals. *Journal of Verbal Learning and Verbal Behaviour*, 11, 455-464.
- Tirre, W. C., & Peña, C. M. (1992). Investigation of functional working memory in the Reading Span test. *Journal of Educational Psychology*, 84, 462-472.
- Tulving E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, 26, 1-12.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.
- Tzeng, O. J. L. (1973). Positive recency effects in delayed free recall. *Journal of Verbal Learning and Verbal Behaviour*, 12, 436-439.
- Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behaviour*, 23, 151-161.
- Van der Linden, M., Coyette, F., & Seron, X. (1992). Selective impairment of the 'central executive' component of working memory: A single case study. *Cognitive Neuropsychology*, 9, 301-326.
- Verhaegen, P., Marcoen, A., & Goossens, L. (1993). Facts and fiction about memory aging: A quantitative integration of research findings. *Journals of Gerontology*, 48, P157-P171.
- Waldrop, M. M. (1987). The workings of working memory. *Science*, 237, 1564-1567.
- Watkins, M. J. (1977). The intricacy of memory span. *Memory & Cognition*, 5, 529-534.
- Watkins, M. J. & Peynircioglu, Z. F. (1983). Three recency effects at the same time. *Journal of Verbal Learning and Verbal Behaviour*, 22, 375-384.
- Watkins, M. J., Watkins, O. C., & Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning & Verbal Behaviour*, 13, 430-477.
- Watzlawick, P., Weakland, J., & Fisch, R. (1974). *Change: Principles of problem formation and problem resolution*. New York: W. Norton.
- Waugh, N. C., & Norman, D. A. (1965). Primary Memory. *Psychological Review*, 72, 89-104.

-
- Weed, K., Ray, E. B., & Day, J. (1990). Metamemory and attributions as mediators of strategy use and recall. *Journal of Educational Psychology, 82*, 849-855.
- Weschler, D. (1981). *Weschler Adult Intelligence Scale Revised*. New York: The Psychological Corporation.
- Whitney, P., Ritchie, B. G., & Clark, M. B. (1991). Working-memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes, 14*, 133-145.
- Wickelgren, W. A. (1965). Short-term memory for phonetically similar lists. *American Journal of Psychology, 78*, 567-574.
- Wimmer, H., & Tornquist, K. (1980). The role of metamemory and metamemory activation in the development of mnemonic performance. *International Journal of Behavioural Development, 3*, 71-81.
- Wingfield, A., Stine, E. A. L., Lahar, C. J., & Aberdeen, J. S. (1988). Does the capacity of working memory change with age? *Experimental Aging Research, 14*, 103-107.

APPENDICES

Appendix 1: Experimental Stimuli Used in the Present Dissertation.

Table A 1. The stimuli from Experiments 1 to 4. In Experiments 1 and 2, stimuli were presented concurrently with the sentence; in Experiments 3 and 4 stimuli were words only. Preceding each sentence is a letter to indicate if a sentence was correct "C" or muddled "M". In Experiments 1 to 4, all stimuli were presented only once; that is, stimuli were drawn from a large pool with replacement.

Sentence		Target item
Phonologically similar items		
C	A sailor might use a navigation	chart
M	Cats are traditionally used to pull a	cart
C	Something extra is a spare	part
C	Paper can be used to make a	dart
M	Jogging is a type of	art
M	A person can easily survive without a	heart
C	To be safe is to be out of	harm
M	One never sees animals on a	farm
M	A book is a type of magical	charm
C	Your hand is at the end of your	arm
C	People are warned of danger by an	alarm
M	Oak trees are a type of	palm
M	On a perfectly calm day one might feel a	breeze
M	Margarine is a type of	cheese
C	To open a lock, one might need a set of	keys
C	A person's legs bend at the	knees
C	Small round green vegetables might be called	peas
M	Land-lubbers are often found sailing the high	seas
C	A pain in the body is called an	ache
M	Gravel is one ingredient of a	cake
M	Water can be moved with a	rake
C	Tomato plants might grow up a	stake
C	In an earthquake, loose objects are likely to	shake
M	A join is the same as a	break
C	A large furry animal could be called a	bear
M	A table is the same as a	chair
C	Candy-floss and merry-go-arounds are found at a	fair
M	A spoon is used to shave off unwanted	hair
C	A portion of something is also called a	share
M	A rip is different to a	tear

M	A cat is a type of	bug
M	One might eat a	rug
M	One can have a bath in a	mug
C	A pot that holds milk is called a	jug
C	Another name for a wheel-nut is a	lug
C	Heroin is a dangerous	drug

Phonologically dissimilar items

C	A jacket, waistcoat, and trousers make up a	suit
C	Yelling and screaming can be called making a	fuss
M	To be very quiet is to	cheer
C	Sickness is the opposite of good	health
C	Another name for a smile is a	grin
M	When a flower dies it is said to	bloom

M	Water is never stored in a	jar
M	Four children can be called	twins
M	The squeak of a mouse is like a lions	growl
C	A Kung-fu expert might use a flying	kick
C	Politicians are elected by a	vote
M	The colour of something is called its	smell

M	Sunshine is the same as	shade
M	Spiders live inside a	shell
C	A label on something is called a	tag
C	A candle can be lit with a	match
M	Scratch is the opposite of	mark
C	A long thin string could be called a	thread

M	Spaghetti is made from dried	worms
C	Before falling off a cliff you first come to the	edge
C	Beds are made with blankets and	sheets
C	Another name for a taxi is a	cab
M	Someone who is very old is called a	kid
M	Rocks are as soft as	straw

C	People's lungs are in their	chest
M	People usually eat off a	bunk
M	It never rains when the sky is full of	clouds
C	A track is a type of	path
C	A tool to arrange hair with is called a	comb
M	It is easy to find someone in a	crowd

C	A view of the countryside can be called a	scene
M	Giraffes have long	trunks
M	Fancy is the same as	plain
C	All sentences must contain a	verb
C	Actors and actresses can act on a	stage
M	Cats have	wings

C	Cats and dogs both have	claws
M	An innocent mistake is the same as a deliberate	trick
M	An increase in prices is called a	sale
C	Dust can be swept away with a	broom
C	Cowboys in the movies sometimes have an Indian	scout
M	An automobile is a type of	yacht
C	Food is served on a	plate
C	Feet are kept warm by	socks
M	A very quiet voice is called a	shout
M	A truism is the same as a	lie
M	A square is more round than a	circle
C	Games played for fun can be called	sport
M	A sandal is heavier than a work	boot
C	Houses can be built in	brick
C	Hot water turns into	steam
M	A lake is smaller than a	tub
M	Good workers always blame their	tools
C	Joined steel rings form a	chain
M	A person found not-guilty would have to pay a	fine
C	Kilograms are a unit of	weight
M	A one word answer could be called a	speech
M	A mansion is smaller than a	hut
C	Where a helicopter lands can be called a	pad
C	Water is moved with a	pump
C	The night before Christmas is called Christmas	eve
M	Darkness comes from a	lamp
C	Something for nothing is a	gift
M	A hill is the same as a	dip
M	A hammer is a type of	food
M	A hairline crack is bigger than a large	gap
C	The liquid from an orange is its	juice
C	Telephones used to all be connected by	wire
M	A flag-pole is smaller than a walking	cane
M	A female cat is called a	bull
C	Smokers use cigarettes, cigars, or a	pipe

Two syllable items

C	A competition is the same as a	contest
C	A car is usually parked in a	garage
M	Work is always the same as a	hobby
C	A field is the same as a	paddock
M	Three people form a	couple
M	The Wright brothers invented the flying	saucer
C	Grandma's husband is usually called	grandpa
C	Food is digested in the	stomach

M	Nastiness is the same as	kindness
M	Milk bottles are made of	leather
M	Cars usually drive on the	footpath
C	A house can be warmed with a	heater
C	Wine can be kept in a	cellar
C	Where someone lives is their home	address
M	At mid-day one can watch the	sunset
M	An elephant is smaller than a	camel
M	A green and square object is called a	snowball
C	A person rowing a boat would use a	paddles
M	Punishment is the same as	reward
C	A long walkway can be called a	trail
C	A school teacher might write a	report
M	Auckland and Wellington are within easy walking	distance
M	A cube is the same as a round	marble
C	Airplanes land at an	airport
C	Buying groceries is called	shopping
C	Attractiveness is another word for	beauty
M	A teacher is the same as a	pupil
M	A bag of worthless stones is called	treasure
M	A flea is larger than a	tiger
C	A group of workers can form a	union
M	The bottom of a lake is called the	surface
M	An ocean liner is smaller than a	canoe
C	A popular movie snack is	popcorn
C	A person who lives next door is called a	neighbour
C	Nails are struck with a	hammer
M	Animals kept in cages have a lot of	freedom

Table A.2. The stimulus words for Experiment 5 for the Fixed-pool condition. Each pool was sampled from with replacement, with the proviso that no word occurred more than once per trial, until 6 stimuli were found. The phonologically similar stimuli were taken from Brookes and Watkins (1990), with the phonologically dissimilar stimuli being taken from Elley (1975) and with a frequency of 5.

Phonological Similarity	
Phonologically similar	Phonologically dissimilar
bug	claws
rug	trick
drug	sale
tug	broom
hug	scout
dug	plate
plug	socks
slug	food
mug	lie
jug	sport

Table A.3. The stimulus words for Experiment 5 for the Non-replacement pool condition. All words were selected from Elley (1975) with a frequency range of 5 to 6.

Phonological Similarity			
Phonologically similar		Phonologically dissimilar	
bone	chart	claws	fuss
cone	cart	trick	cheer
loan	part	sale	health
phone	dart	broom	grin
throne	art	scout	bloom
tone	heart	yacht	jar
land	harm	crown	suit
band	farm	pipe	twins
stand	charm	oak	growl
grand	arm	pear	kick
bland	alarm	trail	vote
hand	palm	pine	smell
chew	breeze	snake	shade
blue	cheese	pump	birth
crew	keys	fit	tag
sue	knees	tiger	match
shoe	peas	flour	mark
glue	seas	route	thread
bleed	ache	seam	chest
weed	cake	shell	bunk
deed	rake	sort	clouds
speed	stake	stock	path
seed	shake	tools	comb
feed	break	weight	crowd
friend	bear	exam	worms
bend	chair	gap	edge
spend	fair	type	sheets
tend	hair	wake	cab
mend	share	grape	kid
lend	tear	tire	straw
bill	bug	load	scene
chill	rug	smell	trunks
grill	mug	wire	plain
spill	jug	worm	verb
skill	lug	deal	stage
pill	drug	aim	wings

Table A.4. The stimuli for Experiment 6 showing the 1 and 3-syllable words. All words were selected from Elley (1975) with frequency 5 to 6.

Word Length			
1 syllable		3 syllable	
vote	claws	bulldozer	basketball
plain	trick	century	Capital
smell	sale	decimal	period
share	broom	tomato	settlement
verb	scout	energy	history
breeze	yacht	factory	medicine
chart	plate	violin	exercise
fuss	socks	committee	December
palm	shout	scientist	General
jar	lie	magazine	canary
bloom	circle	million	Eskimos
ache	sport	junior	enemy
charm	boot	nobody	handkerchief
art	brick	officer	Indian
health	steam	Canada	attention
twins	tub	hamburger	audience
growl	tools	industry	chocolate
kick	chain	pullover	computer
rake	fine	orchestra	article
shade	weight	paragraph	barrier
tag	speech	theatre	carpenter
match	hut	gathering	butterfly
thread	pad	plantation	carnival
mark	pump	avenue	October
chest	wake	period	November
clouds	eve	position	coconut
path	lamp	referees	policeman
straw	gift	average	assembly
crowd	dip	caravan	principal
edge	food	protection	agreement
worms	gap	tobacco	citizen
sheets	juice	typewriter	September
stage	wire	scenery	accident
wings	cane	groceries	addition
kid	bull	happiness	adventure
rug	pipe	calendar	astronaut

Table A 5. The stimuli for Experiment 7. Each pool was sampled from with replacement, with the proviso that no word occurred more than once per trial, until 6 stimuli were found. The phonologically similar stimuli were taken from Brookes and Watkins (1990), with the control and 3-syllable stimuli being taken from Elley (1975) with a frequency of 5.

Control	Phonologically similar	3-syllable
sock	bug	basketball
trick	rug	settlement
sale	drug	chocolate
broom	hug	adventure
chain	tug	carpenter
fine	dug	magazine
hut	plug	industry
pad	slug	committee
dip	mug	happiness
lie	jug	protection

Table A.6. The sentence structure for presenting items in the complex-span task in Experiment 7. The same structure was used for each trial with whether the sentence was correct or muddled and the target word being varied across trials.

The letter '*' is in the word *
 Letter '*' comes after '*' in the word *
 Letter '*' comes before '*' in the word *
 There are 5 letters in the word *
 There are 2 vowels in the word *
 TThere are 2 consonants in the word *

Appendix 2: A Signal Detection Analysis of the Sensitivity and Bias of Participants' Responding on the Sentence Verification Task

Sensitivity and bias for Experiment 1

It could be argued that, in Experiment 1, that differences between stimuli on any measures of dependent variables were a function of a person's ability to decide whether a sentence was correct or muddled. This is an issue addressed by the literature on signal detection theory (e.g., Grier, 1971; MacMillan & Creelman, 1991; Pastore & Scheirer, 1974). To overcome any such criticism, each condition was analyzed for differences in sensitivity and bias estimates (Pastore & Scheirer, 1974; Table A.7).

Table A.7 A summary of participants' correct verifications of correct sentences ($p("c"|C)$) and incorrect verifications of muddled sentences (false alarms; $p("c"|M)$), of the z -values for each probability of a hit or false alarm ($z(H)$, $z(F)$), a sensitivity index ($d' = z(H) - z(F)$), and a bias index ($c = -.5 [z(h) + z(F)]$) for both articulatory suppression and phonological similarity conditions in Experiment 1. d' values of more than 3 and bias estimates of less than ± 0.5 indicate high sensitivity and low bias respectively (MacMillan & Creelman, 1991).

	Hit rate	False alarm rate	Sensitivity (d')	Bias (c)
Control				
Similar	.99 (.03)	.07 (.09)	3.98 (0.67)	-.23 (.42)
Dissimilar	.98 (.05)	.07 (.09)	3.87 (0.82)	-.22 (.37)
Articulatory suppression				
Similar	.92 (.12)	.10 (.09)	3.20 (1.01)	-.12 (.49)
Dissimilar	.97 (.06)	.06 (.07)	3.80 (0.88)	-.12 (.35)

The mean hit and false-alarm rates were calculated across all conditions. From the mean hit and false-alarm rates, the overall d' and c indices were calculated. It was apparent from these overall indices of sensitivity and bias, respectively, that that participants were extremely able in differentiating as to whether a sentence was correct or muddled ($d' = 3.8$ over all conditions and persons; Table A.7) and that participants were only very slightly more inclined to report sentences as correct than muddled ($c = -.14$; Table A.7), so slight that for practical purposes participants were

unbiased. What bias existed is most likely explained as accidental incorrect key-presses and ambiguous verifications, both of which were reported by some participants reported. Thus when analyzing verification times, there is no basis to expect confounding of either participant sensitivity or bias (c.f., Tirre & Peña, 1992).

It was also apparent in Experiment 2 that all participants in every condition were extremely able in differentiating as to whether a sentence was correct or muddled ($d' = 3.8$ overall conditions and persons; Table A.8) and were only very slightly more inclined to report sentences as correct than muddled ($c = -.10$; Table A.8).

Table A.8. A summary table of participants' correct verifications of correct sentences ($p("c"|C)$) and incorrect verifications of muddled sentences (false alarms; $p("c"|M)$), of the z -values for each probability of a hit or false alarm ($z(H)$, $z(F)$), a sensitivity index ($d' = z(H) - z(F)$), and a bias index ($c = -.5 [z(h) + z(F)]$) for both suppression and word length conditions in Experiment 2. d' values of more than 3 and bias estimates of less than ± 0.5 indicate high sensitivity and low bias respectively (MacMillan & Creelman, 1991).

	Hit rate	False alarm rate	Sensitivity (d')	Bias (c)
Control				
1-syllable	.96 (.06)	.05 (.06)	3.82 (.82)	-.07 (.38)
2-syllable	.98 (.04)	.03 (.05)	4.17 (.67)	-.03 (.32)
Articulatory suppression				
1-syllable	.97 (.05)	.05 (.08)	3.98 (.91)	-.09 (.31)
2-syllable	.99 (.03)	.02 (.05)	4.32 (.63)	-.07 (.23)

Appendix 3: The General Information Sheet for Experiments 1 to 7

Experiment 'X' - Working Memory span

My name is Llewelyn ("Llew") Richards-Ward. I can be contacted through the Psychology Department office at Massey University. Just phone 06-356-9069 and ask for the Psychology Office.

I am doing this study as part of my doctoral thesis in psychology. You may notice from the title that **my research involves memory**. What this study is attempting to do is to examine whether the Reading Span task that you will do can be linked to one theory of working memory. There are no tricks and **this study is in no way a 'test' or clinical assessment of your memory**.

To do this study you will need to attend the Psychology Department at a pre-arranged time. I will provide anyone with a map who requires one. After you arrive and we meet each other, I will ask you to complete an informed consent sheet and a screening questionnaire. Then you will have a **task on a computer** explained to you. Please do not be put off by the computer. If you can use a telephone keypad, you can do this task! You will be given a few practice trials to help you feel comfortable about the task.

All of this will take about **1 hour** including rest breaks.

There are also some things that you can expect of me. If you take part in the study, you have the right to:

- **refuse to answer** any particular question.
- **withdraw** from the study at any time.
- **ask questions** about the study that may occur to you during your participation.
- **confidentiality** of the information that you provide. It will not be possible to identify you in any reports that are prepared from the study.
- **a summary** of the findings from the study when it is concluded.

Appendix 4: The Participant Consent Sheet for Experiments 1 to 7

Working Memory and Memory Self-efficacy studies.

Llewelyn Richards-Ward

INFORMED CONSENT

Name: _____

Address: _____

Phone: _____

I have read the Information sheet for this study and have had the details of the study explained to me. My questions about the study have been answered to my satisfaction, and I understand that I may ask further questions at any time.

I also understand that I am free to withdraw from the study at any time, or to decline to answer any particular questions in the study. I agree to provide information to the researcher on the understanding that it is completely confidential.

I wish to participate in the study under the conditions set out in the information sheet.

Signed: _____

Date: __/__/__.



Appendix 5. F-values and errors terms for Experiment 7

Table A.5.1. Planned comparisons of control words against phonologically similar words.

	F-Value	M.S.E.
<i>Recall in the correct serial position</i>		
1. Tasktype	2.33	1.55
2. Phonological similarity	34.32***	0.57
1 x 2	1.07	0.49
1. Tasktype	4.76*	2.31
2. Word length	32.10***	0.41
1 x 2	1.38	0.35
<i>Recall in any serial position</i>		
1. Tasktype	1.66	1.06
2. Phonological similarity	16.52***	0.25
1 x 2	0.54	0.16
1. Tasktype	3.94	1.43
2. Word length	30.83***	0.24
1 x 2	9.85**	0.18
<i>Order errors</i>		
1. Tasktype	1.08	0.31
2. Phonological similarity	22.81***	0.25
1 x 2	4.14	0.25
1. Tasktype	2.03	0.44
2. Word length	2.95	0.27
1 x 2	1.96	0.21
1. Tasktype	0.01	0.41
2. Word length vs. Phonological similarity	9.35**	0.25
1 x 2	0.70	0.19
<i>Viewing time</i>		
1. Tasktype	2.33	2 765 611
2. Phonological similarity	3.42	65 573
1 x 2	0.54	59 670
1. Tasktype	3.60	3 740 884
2. Word length	33.88***	110 736
1 x 2	17.31***	99 201
<i>Post-estimates</i>		
1. Tasktype	72.10***	0.60
2. Phonological similarity	20.83***	0.38
1 x 2	4.17*	0.26
1. Tasktype	76.81***	0.69
2. Word length	41.50***	0.30
1 x 2	0.36	0.29
<i>Pre-estimates</i>		
1. Tasktype	92.11***	0.18
2. Phonological similarity	9.83**	0.10
1 x 2	0.77	0.07
1. Tasktype	107.24***	0.19
2. Word length	43.73***	0.17
1 x 2	6.35*	0.07

Table A.5.1 cont.

<i>Articulation rate</i>		
1. Wordtype	5.77†**	4 960 251
2. List	0.12†	10 594
		628
1 x 2	4.92†*	5 981 802
1. Phonological similarity	12.34***	0.04
2. Word length	0.23	0.13

Note: For all analyses, d.f.=1,31 except where indicated († d.f.= 2, 62).

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