The relationship between storage and processing in working memory

Conlin, Juliet Ann

Conlin, Juliet Ann (2005) *The relationship between storage and processing in working memory*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/1787/

The Relationship between Storage and Processing in Working Memory

Juliet Ann Conlin

A thesis submitted to the University of Durham for the degree of

Doctor of Philosophy

2005

University of Durham
Department of Applied Psychology
Queens Campus
Stockton-on-Tees, UK

Juliet Ann Conlin

Thesis title: The Relationship between Storage and Processing in Working Memory

Abstract

This thesis reports a series of eight experiments that investigate the detailed nature of the factors underpinning working memory performance in children and adults. Experiments 1 to 3 examined the role of resource-sharing and intrinsic memory demands in complex span performance in 7- and 9-year-olds. The results do not support a resource-sharing explanation and are consistent with the view that complex memory span performance is disrupted by processing activities that divert attentional resource from storage. Experiments 4 and 5 investigated the impact of the similarity of processing and storage stimuli on span performance of 7- to 9-year-old children. The data provide evidence for performance-related decrements under circumstances of stimulus similarity. Experiments 6 to 8 investigated the impact of the lexical status of memory and processing stimuli on children's and adults' complex memory performance with the aim of exploring possible mechanisms of interference in working memory. In 9- and 10-year old children and adults, word recall was markedly impaired by monitoring words compared with nonwords. A converse disturbance of nonword recall by nonword monitoring was consistently found for adults, but was either absent or less marked across experiments in the child groups. Overall, the data identify four main factors that mediate complex span performance in children and adults: task duration, attentional resources, task-switching efficiency, and interference between processing and storage stimuli. The findings are discussed in terms of existing theoretical models of working memory.

Table of contents

Appendix I

Appendix II

List of tables

List of Figures

Figure 1:

Figure 2

Declaration

None of the data or material contained in this thesis has been submitted for previous or simultaneous consideration for a degree in this or any other university.

Chapter 2 is reported in:

Conlin, J.A., Gathercole, S.E., & Adams, J.W. (2005). Children's working memory: Investigating performance limitations in complex span tasks. *Journal of Experimental Child Psychology, 90*, 303-317.

Chapter 3 is reported in:

Conlin, J.A., Gathercole, S.E., & Adams, J.W. (in press). Stimulus similarity decrements in children's working memory span. *Quarterly Journal of Experimental Psychology.*

Chapter 4 has been submitted to:

Conlin, J.A., & Gathercole, S.E. (2005). Lexicality and interference in working memory. *Journal of Memory and Language.*

Note: The lexical cue discrimination hypothesis discussed in Chapter 4 was developed in collaboration between the thesis author and her supervisor, Professor Susan E. Gathercole

Statement of copyright

The copyright of this thesis rests with the author. No quotation from it should be published in any format, including electronic and the Internet, without the author's prior written consent. All information derived from this thesis must be acknowledged appropriately.

Acknowledgements

I would like to thank my supervisors Sue Gathercole and John Adams for their invaluable encouragement and advice, and especially to Sue for her confidence in my ability to complete this project successfully.

I am also grateful to all those who participated in the experiments, without whose cooperation this thesis would not be possible.

Love and special thanks to my husband Michael, for his thoughtful suggestions and unwavering support. Thanks also to my children, Jake and Fay, for giving me peace and quiet when it mattered, and for their interest and enthusiasm.

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

Chapter 1

General Introduction

Much of everyday cognition is dependent on the ability to retain information temporarily while mentally engaging in its transformation. An obvious example of this process is mental arithmetic: one must bring to mind the procedure with which to solve the problem, retain the numbers in immediate memory while calculating the answer, perhaps even with the requirement to store interim solutions. Thus, two features that characterise such "working memory" activities are those of information storage and processing.

There is comprehensive evidence that performance on tasks that incorporate concurrent storage and processing activities – unlike measures of simple recall ability – is linked in both children and adults with key cognitive skills such as language comprehension (e.g., Daneman & Merikle, 1996), reasoning (e.g., Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002), general fluid intelligence (e.g., Engle, Tuholski, Laughlin, & Conway, 1999), and with learning abilities in the areas of both literacy (e.g., Swanson, Ashbaker, & Lee, 1996) and mathematics (e.g., Geary, Hoard, Byrd-Craven, & DeSoto, 2004), although the underlying cognitive processes that support performance on such tasks remain open to debate. This thesis is concerned with the mechanisms that underpin the relationship between storage and processing activities in working memory, which – though under extensive study in recent years – are still controversial and have yet to be fully specified.

In an introduction to the experiments reported in this thesis, this chapter will review current models of working memory, describe the measures commonly used to assess working memory performance in children and adults, and evaluate the theories advanced to account for limitations on such tasks. The final section outlines the major points addressed in this thesis.

1.1. Working memory

1.1.1. A multi-component model

In an attempt to formalise and account for the huge range of experimental findings on short-term memory phenomena that had been collected over the past century, Baddeley and Hitch (1974) proposed a model that incorporated a working memory akin to the central processor of Craik and Lockhart (1972), and an articulatory rehearsal loop similar to the echoic store proposed by Atkinson and Shiffrin (1968). In addition, a further component was hypothesised to handle the storage and manipulation of visual and spatial material. The model was termed the "Working Memory Model", and has since its introduction come to form a central construct in cognitive psychology.

1.1.2. The central executive

According to the model, which has been developed subsequently by Baddeley and colleagues (Baddeley, 1996, 2000; Baddeley & Logie, 1999), working memory reflects multiple resources associated with a controlling attentional system that supervises and coordinates a number of distinct capacity-limited sub-systems. The attentional component is termed the central executive, which is responsible for the control of encoding and retrieval strategies, high-level processing activities, the switching of attention, and the coordination of activities within working memory.

1.1.3. The phonological loop

Other components of the working memory model include two modality-specific slave systems: the phonological loop and the visuo-spatial sketchpad. The phonological loop is a system capable of maintaining and refreshing verbal information. As the most comprehensively researched of the model's components, evidence suggests that it is sub-divided into two components: a passive short-term store and an active sub-vocal rehearsal system (e.g. Baddeley, Lewis, & Vallar, 1984). The contents of the passive store are subject to decay, but can be refreshed and maintained by subvocal rehearsal. Evidence for the characteristics of the phonological loop comes from studies that have found disruptive effects of phonological similarity (e.g. Baddeley, 1966), irrelevant speech (e.g. Salamé & Baddeley, 1982), word length, and articulatory suppression (e.g. Baddeley, Thomson, & Buchanan, 1975). The phonological similarity effect describes the decline in recall for words that sound similar, and is thought to occur because items contained within the passive short-term store will become confused when they are phonologically similar to one another. The effect of irrelevant speech arises from the fact that heard speech directly accesses the phonological store, thereby disrupting its current contents. When participants are required to suppress articulation by repeating aloud an irrelevant phoneme such as "*the, the, the*", temporary memory for sequences of verbal items is disrupted. The working memory model accounts for this effect by suggesting that the technique blocks the use of the active subvocal rehearsal mechanism, thereby undermining an important facility for retention (Baddeley et al., 1984).

Working memory research has benefited from recent advances in technology, which allow explicit tests of theories of working memory. In general, cognitive neuroscience research suggests a major role for the prefrontal cortex in working memory. For example, a positron emission tomography study using an *n*-back task showed localised rehearsal in the frontal speech areas in the brain (Awh, Jonides, Smith, Schumacher, Koeppe, & Katz, 1996). The authors concluded that these frontal regions used in spoken language are recruited for the purposes of maintaining verbal information active in working memory, and are distinct from passive short-term storage. This provides neat support for Baddeley et al.'s (1984) notion of the phonological loop comprising an active sub-vocal rehearsal system and a distinct passive short-term store.

1.1.4. The visuo-spatial sketchpad

The visuo-spatial sketchpad component of working memory has not been explored to the same extent as the phonological loop, although there has been an increase in research interest in recent years (e.g. Bruyer & Scailquin, 1998; Pearson, Logie, & Gilhooly, 1999; Duff & Logie, 1999). Evidence points to a fractionation of this component into a passive visual cache, responsible for the retention of visual patterns, and an inner scribe, an active spatially-based rehearsal mechanism (Logie, 1995). Several studies have demonstrated a disruptive effect of concurrent movement on the retention of spatial patterns (e.g. Smyth & Pendelton, 1989; Logie, Zucco, & Baddeley, 1990), and the viewing of irrelevant, changing visual material can disrupt the retention of visual information (e.g. Quinn & McConnell, 1996). Further evidence suggests – analogous to the phonological similarity effect – that confusions can arise in memory for visually similar material (e.g. Logie, Della Sala, Wynn, & Baddeley, 2000).

1.1.5. The episodic buffer

Originally hypothesised to comprise only three components, the Working Memory model was recently modified to include a fourth: the episodic buffer (Baddeley, 2000). This component was added to account for an increasing body of data that provided considerable problems for the original conceptualisation as a tripartite structure. For example, brain damaged patient P.V. demonstrated a word span of one, but a sentence span of five words (Vallar & Baddeley, 1984), indicating the existence of a system other than the phonological loop for storing verbal information. Logie et al. (2000) report evidence that visual and phonological factors exert a concurrent influence on the recall of verbal information. Hence, the recently proposed episodic buffer is thought to be responsible for the temporary storage of multi-modal information and for integrating representations both within subsystems of working memory and across the cognitive system more generally. In addition, the buffer is hypothesized to serve as an interface between working memory and long-term memory, facilitating input to and retrieval from long-term memory. This system has been used to explain marked individual differences in the immediate recall of prose by amnesic patients (Baddeley & Wilson, 2002), and could also account for the finding that twice as many words can be recalled in the correct order when they form a meaningful sequence (Brener, 1940; Baddeley, Vallar, & Wilson, 1987). However, empirical evidence to support the notion of this recent addition to the working memory model remains scarce.

1.1.6. Semantic short-term memory

Neuropsychological studies of brain-damaged patients have led some researchers to suggest that the storage of verbal information is supported not only by a phonological short-term store, but also by a semantic short-term memory (Hanten & Martin, 2000; Martin & Freedman, 2001), thus questioning the structure of current models of working memory (e.g., Baddeley, 1986) that assume that verbal storage is handled exclusively by the phonological loop. Martin, Shelton, and Yaffee (1994) tested two brain-damaged patients on short-term memory tasks for phonological and semantic codes. The phonological tasks consisted of a traditional digit span task and a digit matching span task. The semantic short-term memory tasks consisted of a test for memory for words over nonwords; the other task involved participants listening to a list of words and then judging either a) whether a probe word rhymed with one of the list words, or b) whether the probe word belonged to the same semantic category as one of the list words. The patients tested displayed the following memory deficits: E.A. demonstrated a greater phonological than semantic short-term memory deficit; A.B. showed the opposite pattern, in that his performance on tasks that required phonological activity was better than on semantic tasks. E.A.'s performance on the probe recognition task was significantly lower when the judgement involved rhyme detection than semantic category detection, suggesting a phonological impairment. A.B., in contrast, appeared to have a relatively intact phonological short-term memory, as demonstrated by a normal word length effect (i.e., superior recall of short over long words), and a normal modality effect (i.e., superior recall of words following auditory than visual presentation). In addition, A.B. did not show a lexicality effect, that is, better recall of words over nonwords, indicating an impairment of his semantic short-term memory system.

The separability of phonological and semantic short-term memory has received support from studies with healthy adults (Haarman, Davelaar, & Usher, 2003), indicating that working memory performance may be additionally constrained by semantic, non-phonological processes in verbal short-term memory. Taken together, the evidence from neuropsychology and neuroimaging studies of working memory clearly indicates that a cognitive neuroscience approach can offer a useful supplementary means of assessing the structure of working memory with regard to its biological implementation.

1.1.7. Limitations of executive functioning

The working memory model offers a parsimonious account of a large body of data, ranging from language acquisition (e.g., Gathercole & Baddeley, 1989) to visual imagery (Logie, 1995), although critics have pointed to the underspecification of some of the components, most notably the central executive (e.g. Towse & Houston-Price, 2001), in the sense that there is "*... a temptation to invoke the central executive to explain any aspect of cognitive data which cannot be attributed to the phonological loop or visuo-spatial* *sketchpad*' (Andrade, 2001, pp. 285). Specifically, the exact nature and limitations of executive functioning in working memory – such as selective attention, dual-task coordination, and strategy selection – remain contentious in current working memory research. An understanding of these functions is crucial in illuminating the processes at work during complex cognitive activities. For example, what are the mechanisms underlying the coordination of executive processing and temporary storage during a task such as reading comprehension or mental arithmetic? What limits performance on cognitive tasks: short-term storage capacity, processing speed, interference among representations, or complexity of the processing operations? Under which circumstances, if any, do storage and processing compete for resources?

1.2. Working memory and attention

1.2.1. Focus of attention

One candidate for explaining limitations in working memory is that of a limited-capacity focus of attention. This alternative view of working memory was advanced by Cowan (1988; 1993; 1995), who made a distinction between short-term memory and working memory. According to this account, working memory capacity is constrained by the amount of information that, once activated above a certain threshold, can be held in the focus of attention. The amount of information (including sensory, phonological and semantic material) that can be activated is unlimited, but it is subject to decay within up to 20 seconds (Cowan, 1984). In contrast, information that is held in the focus of attention does not decay, but there is a capacity limit on how much information can be in focus. Maintaining the information in focus requires controlled, limited-capacity attention, and interference can occur among activated items.

1.2.2. Controlled attention

A related view of working memory is proposed by Engle and colleagues (e.g., Turner & Engle, 1989; Engle, Cantor, & Carullo, 1992; Engle et al., 1999). According to their general capacity model, working memory comprises a store in the form of long-term representations activated above some critical threshold (short-term memory), and a limited-capacity attentional system for maintaining the activation of these representations. This attentional capacity is regarded to be domain-free, and consequently, individual differences in this capacity can be observed across a wide variety of cognitive tasks. Critical to this view is the importance of attention under conditions of interference, during which the maintenance of information in the focus of attention is most difficult. In conditions in which interference is absent, task-relevant information may be retrieved from LTM relatively easily or automatically.

Thus, the focus of attention postulated by Cowan (e.g., 1995) and Engle et al.'s (e.g., 1999) working memory can be considered analogous to Baddeley's (1986) central executive, although these accounts differ from the Baddeley (1986) model in that there are no functionally distinct phonological and visuo-spatial subsystems responsible for short-term storage. Instead, remembering (as opposed to processing) in working memory is thought to depend on domain-specific skills that facilitate storage, such as chunking and rehearsal.

1.3. Measures of working memory

1.3.1. Complex span tasks

As a construct widely regarded to comprise storage and processing functions, working memory is most appropriately measured by tasks designed to reflect the simultaneous retention and manipulation of information. One of the most commonly used tasks developed to measure working memory capacity is the complex span task. Unlike simple span tasks, which assess the maximum amount of information that can be stored for a short period in memory, complex span tasks require participants to retain information in the face of ongoing mental activity. Hence, whereas simple span tasks typically involve the immediate serial recall of lists of information, the complex span task paradigm is characterised by the alternate presentation of to-be-remembered stimuli, such as words, digits, or letters, and stimuli that require some form of mental processing, such as sentence comprehension or mental arithmetic (for a review, see Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, in press).

There is abundant evidence to suggest that simple span measures are not suitable measures of working memory capacity (e.g., Daneman & Carpenter, 1980; Engle et al, 1999; Conway, Cowan, Bunting, Therriault, & Minkoff, 2002). In the word span task, for example, participants are presented with a list of items and must repeat the sequence in serial order. Span is taken as the maximum number of words that can be recalled accurately. As such, this span task involves relatively simple processes such as rehearsal and retrieval of common lexical items (e.g., Daneman & Carpenter, 1980). Word span's utility as a working memory measure is questioned by the fact that it failed, for example, to distinguish between good and poor readers who were classified on the basis of a standardised reading comprehension test (e.g., Guyer & Friedman, 1975), a test that is assumed to rely heavily on working memory. In addition, simple span measures correlate poorly with complex cognitive tasks, both in adults (e.g., Turner & Engle, 1989) and children (e.g., Leather & Henry, 1994), although several recent studies provide exceptions to this (e.g. Towse & Houston-Price, 2001; Bayliss, Jarrold, Baddeley, & Gunn, 2005).

1.3.2. The reading span task

Daneman and Carpenter (1980, Experiment 1) were among the first researchers to utilise a complex span task to examine the relationship between reading comprehension and verbal working memory capacity. In this task, which they named the reading span task, participants were required to read a set of sentences aloud at their own pace, while simultaneously trying to remember the last word of each sentence. An example sentence is: *The taxi turned up Michigan Avenue where they had a clear view of the lake.* The span task contained sets of 2, 3, 4, 5, and 6 sentences, and participants were presented with increasingly longer sets of sentences until they failed to recall, in serial order, the sentence-final words from all three sets at a particular level. Testing was terminated at this point. The level at which a participant correctly recalled two out of three sets was taken as a measure of that person's reading span. So, for example, if a participant correctly recalled the sentence-final words from two sets of three sentences, the number of sentences in a set would increase by one. If the participant failed to recall the sentence-final words in more than one of the three sets of four sentences, their reading span would be three. The number of words recalled in the face of ongoing processing was interpreted by Daneman and Carpenter (1980) as representing the residual storage capacity of working memory. In other words, reading span is a measure of the working memory capacity that is not allocated to the processing portion of the task. The researchers underlined this interpretation with the finding that performance on the reading span task was a better predictor of reading comprehension than a simple word span task administered to the same participants.

1.3.3. Operation and counting span

Over the past 25 years, this working memory task has since been developed and extended to include other forms of storage and processing requirements. For example, in order to demonstrate that the relationship between reading comprehension and complex span task performance is not specific to reading ability – as originally proposed by Daneman and Carpenter (1980) – Turner and Engle (1989, Experiment 1) replaced the sentences in the task with mathematical operation strings. In this task, termed 'operation span task', participants were required to verify a visually presented solution to the mathematical equation, which was correct on half of the trials, and recall, in order, unrelated words that were presented immediately following the participants' verbal verification of the operation task. An example of the type of operation-word string used is: "[(9/3) – 2 = 11] house". Turner and Engle (1989) found that performance on this task led to correlations with reading comprehension similar to those found when the secondary task was reading, and argued on the basis of these results that working memory capacity is independent of the specific nature of the processing component of the task.

Another widely used complex span measure is counting span. Originally developed by Case, Kurland and Goldberg (1982), the counting span task requires participants to count a series of visually presented arrays of shapes, and to remember the count totals for subsequent recall. Due to the relative simplicity of the processing requirement of the task (i.e., counting), this task is often used when testing children, and findings from studies using the counting span task (e.g., Case et al., 1982; Hutton & Towse, 2001) demonstrate that complex span tasks are useful in measuring working memory capacity not only in adults, but also in child populations.

1.3.4. Spatial span

Other aspects of working memory performance, such as spatial thinking, have also been assessed using complex span tasks. Shah and Miyake (1996) developed a spatial span task, in which participants were required to judge whether a set of individually presented letters was normal or mirror-imaged while keeping track of the orientation of the individual letters. The participant had to recall, at the end of the trial, the orientation of each letter in the order in which it had appeared. Importantly, Shah and Miyake found that spatial span correlated with spatial visualisation measures, but not with verbal ability measures, indicating that the nature of the processing activity is crucial in determining the complex span tasks' predictive power.

1.3.5. Reliability of complex span tasks

The abundance of research based on the use of complex span tasks has provided ample evidence of span scores' reliability; that is, complex span tasks produce very similar results from one occasion to another. Reliability analyses across time have found that adults' span scores remain stable (test-retest correlations of approximately .70 to .80) over minutes (e.g. Turley-Ames & Whitfield, 2003) and weeks (e.g. Friedman & Miyake, 2004). Cross-age consistency was observed for children's counting span in three waves of yearly data (Ransdell & Hecht, 2003), and Hitch, Towse, and Hutton (2001) reported test-retest correlations of .71 and .56 for operation span and reading span, respectively, over a period of one year. Internal consistency for complex span task performance was found, for example, by Oberauer, Süß, Schulze, Wilhelm, and Wittmann (2000), who observed Cronbach's alphas of .84 for reading span and .86 for spatial span, indicating that participants' responses were consistent across items within a task.

1.4. Working memory span tasks and complex cognitive abilities

1.4.1. Predictive power

Despite the fact that the debate still continues over what constitutes the 'core' of working memory (e.g. Miyake & Shah, 1999), these tasks remain common research tools in cognitive psychology. One aspect of complex span tasks that continues to generate considerable research interest and theoretical debate is that of the tasks' predictive validity with regard to performance on 'real-word' complex cognitive activities. As mentioned earlier, there is comprehensive evidence that performance on complex span tasks is strongly related to higher level cognitive abilities in both children and adults such as reasoning and reading comprehension (e.g., Engle et al., 1992; Kyllonen & Christal, 1990) and also academic achievement (e.g., Swanson, 1999; Gathercole, Pickering, Knight, & Steadman, 2004).

1.4.2. Language comprehension and complex span

In a comprehensive meta-analysis of data from 77 studies, Daneman and Merikle (1996) came to the conclusion that complex span measures that included a storage and processing component were far better predictors of language comprehension than storage measures alone. This was the case for complex span tasks involving verbal storage and processing ($r = .41$ on global comprehension measures and $r = .52$ on specific comprehension measures) as well as maths storage and processing ($r = .30$ on global comprehension measures and $r = .48$ on specific comprehension measures), compared to verbal storage (.28 and .40 for global and specific measures respectively), and numerical storage (.14 and .30 for global and specific measures respectively). In addition, both reading span and operation span still predict comprehension (albeit to a lesser extent) when individual differences in processing efficiency are statistically controlled (Conway & Engle, 1996; Engle et al., 1992). With research evidence mounting that complex span tasks can predict performance on complex cognitive tasks, the question remains open as to why this is the case. This is not just of theoretical interest, but also in the context of education and development, as complex span performance would appear to be a factor in children's reading and number skill acquisition (Hitch et al., 2001).

1.5. What do complex span tasks measure?

1.5.1. Resource-sharing ability

As described earlier, Daneman and Carpenter (1980) found high correlations between reading span and three measures of reading comprehension: answering fact questions (*r* = .72), pronoun reference questions (*r* = .90), and the Verbal Scholastic Aptitude Test (*r* = .59 in Experiment 1; *r* = .49 in Experiment 2). The researchers argued that reading span performance was linked to individual differences in reading comprehension due to the variability between readers in the efficiency of their reading skills. Thus, good readers utilise less working memory capacity during the processing phase of the task (i.e. reading), leaving a larger portion of working memory available for storage. Daneman and Carpenter proposed therefore that the predictive power of the reading span task was dependent on a specific processing task, namely reading. Consequently, individuals who are good readers and perform well on the reading span task will not necessarily outperform poor readers on a task in which the processing involves a task other than reading.

1.5.2. General capacity hypothesis

However, as described above, Turner and Engle (1989) found that an operation-word span score predicted reading comprehension as well as did a reading span score. On the basis of these findings, Turner and Engle argued that the significant relationship between complex span tasks and reading comprehension was the result of a relatively stable capacity which transcends the specific task. In their view, a complex span task measures the number of items that can be kept active in memory in the absence of mnemonic strategies such as rehearsal or chunking. The underlying assumption of this view is that processing and storage activities compete for a general resource, and that when memory demands increase, there is less capacity available for processing and *vice versa.*

1.5.3. General fluid intelligence and attention

Engle et al. (1999) used structural equation modelling to asses the fit of data from 133 students' performance on a wide range of memory tasks, tests of general intelligence, and standardised academic tests. Engle et al. identified a two-factor model as the best fit of the data, comprising working memory and short-term memory as distinct constructs. Of the two constructs, working memory showed a strong predictive relationship with general fluid intelligence, with no significant association between short-term memory and intelligence. When the variance common to the short-term memory and working memory latent variables was statistically removed, the left-over or residual variance was highly and significantly associated with general fluid intelligence. This residual variance was interpreted by Engle et al. as reflecting a capacity for controlled attention. The authors argued that complex span tasks measure the capacity of short-term memory *plus* the construct controlled attention, and concluded that the component of complex span tasks that is important to complex cognitive functioning is controlled attention. Importantly, this view holds that complex span tasks can predict cognitive performance across a variety of domains, because of the general attentional demands of the tasks, rather than the domain-specific features of the tasks.

1.5.4 Multiple components

Within Baddeley's (1986) model, the storage demands of complex memory span are suggested to depend on appropriate subsystems, with processing supported principally by central executive resources (Baddeley & Logie, 1999; Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002). Again, this view is not incompatible with the Engle et al. (1999) account, if one assumes that controlled attention reflects the properties of the central executive, and that the temporary storage of items (Engle et al.'s short-term memory construct) is handled by the phonological loop (verbal material) or the visuo-spatial sketchpad (visual and spatial material).

1.5.5. Domain-specificity vs. domain-generality

Bayliss, Jarrold, Gunn, and Baddeley (2003) provided a systematic investigation of the constraints underlying individual differences in working memory performance. Specifically, the rationale of the Bayliss et al. study was to assess the extent to which processing and storage abilities are domain-general or domain-specific; that is, whether individual differences in processing efficiency and storage capacity contribute independently to complex span performance. Unitary models of working memory assume that individual differences in storage capacity are related to individual differences in processing efficiency (e.g. Case et al., 1982; Daneman & Carpenter, 1980). Contrary to such models, Bayliss et al. provided evidence that complex span performance was constrained by individual differences in domain-general processing efficiency and domain-specific storage capacity, in both children and adults.

A further important finding from the Bayliss et al. (2003) study was that when the variance associated with the processing and storage activities on the complex span tasks was removed, this residual variance was predictive of children's reading and mathematics performance if the storage domain of the task was verbal, and predictive of adults' reading and mathematics performance regardless of the storage domain of the task. Bayliss et al. argued that this residual variance may well reflect an additional ability in complex span tasks, namely the coordination of processing and storage requirements in working memory. This view is consistent with findings from structural equation modelling studies by Oberauer, Süß, Wilhelm, and Wittmann (2003) and Miyake, Friedman, Emerson, Witzki, and Howerter (2000), who suggest that dual tasking, that is, the coordination of two tasks performed simultaneously, may tap an executive function whose role it is to form new relationships between elements held in working memory.

1.5.6. Simple span and complex cognitive abilities

While the vast majority of studies looking at the predictive power of complex span tasks claim that these tasks tap something simple span tasks do not (e.g., Daneman & Carpenter, 1980; Turner & Engle, 1989), several studies have reported that in children, simple spans *are* significantly associated with cognitive abilities under certain circumstances (e.g., Bayliss et al, 2005; Engle, Carullo, & Collins, 1991; Towse & Houston-Price, 2001; see also LaPointe & Engle, 1990, for evidence that simple word span correlated with reading comprehension in adults). Hutton and Towse (2001) argued that it would be unwise to presume that identical – albeit more efficient – mechanisms underlie both children's *and* adults working memory performance. In their study, the relationship between short-term memory measures and cognitive abilities (here, reading and number) was superior to the complex span scores after processing time and age had been controlled for, indicating the need for caution when using adult data in making inferences regarding children's working memory performance. The following section addresses issues surrounding the development of working memory, specifically with regard to the question of whether developmental changes are best described as qualitative or quantitative.

1.6. Development of working memory

1.6.1. Use of strategies

Almost all measures of short-term memory show a steady increase from the preschool years through to adolescence (e.g. Gathercole, Pickering, Ambridge, & Wearing, 2004; Hulme, Thomson, Muir, & Lawrence, 1984; Siegel, 1994). This may be partly due to an expansion in functional capacity or processing speed, but evidence suggests that this is also due to the development of strategies. For example, the use of strategies that aid recall, such as rehearsal and chunking, has been observed to develop gradually (see Gathercole & Hitch, 1993, for a review), although there is research evidence to suggest that developmental differences in strategy use are less than adequate in accounting for observed developmental differences in memory span (Reyna & Brainerd, 1991).

1.6.2. Changes in knowledge

Changes in strategy use can be seen to be linked to changes in knowledge; for example, knowing which strategies are useful and how to implement them. Chi (1978) found that ten-year old chess experts showed much better immediate recall of a chess board formation than adults who did not play chess, despite the fact that children's digit spans were much lower than those of adults. This certainly suggests that knowledge, or experience, can aid performance in working memory. More specifically, Chi's (1978) data can be interpreted as indicating that the overall capacity of working memory does not increase with age; instead, older people's greater knowledge (acquired through experience) can serve to store more information in more meaningful chunks.

1.6.3. Tripartite model

Studies into the developmental trajectory of the individual components of Baddeley's (1986) working memory model suggest that the structural organisation of working memory remains more or less constant over the childhood years. For example, Gathercole et al. (2004) reported linear increases in performance from 4 years of age to adolescence, with no evidence of consistent developmental changes in the relationship between the central executive and its slave systems. This suggests that children's performance on complex span tasks such as reading span is limited not only by the capacity of the central executive, but is also constrained by the amount of material that can be held in the phonological loop. In addition, these findings indicate that developmental differences in performance on working memory tasks are quantitative rather than qualitative, and provide evidence that the same mechanisms underpin working memory performance across development.

1.6.4. Operational efficiency

According to an influential account advanced by Case et al. (1982), working memory comprises two components: operating space and short-term memory space. Case et al. proposed – in line with Chi's (1978) suggestion – that the development of working memory through to adulthood does not occur through a change in the size of the total processing space (i.e., working memory capacity). However, Case et al. argued, contrary to Chi's view, that the developmental increase cannot be ascribed entirely to changes in mnemonic strategies or efficient chunking, but instead to a decrease in the proportion of this space that must be devoted to cognitive operations. Accordingly, developmental differences arise as a result of changes in operational efficiency. This interpretation is based on the idea of an executive processing space that is allocated to either storage functions or processing functions. Importantly, Case (1985) specifies that "*operating space and short-term memory storage space do not imply two different capacities, ... they imply one capacity that can be flexibly allocated to either of two functions*" (p. 290).

1.7. The relationship between processing and storage

1.7.1. Trade-off between processing and storage resources

An initial account of the cognitive processes underpinning working memory span was that the processing and storage demands of the tasks compete for a limited resource. By this account, increases in processing efficiency result in the availability of additional resources to support storage (e.g., Daneman & Carpenter, 1980). This concept of working memory was used by Case and colleagues as the basis for an account of the developmental increases in working memory span performance across the childhood years (Case, 1985; Case et al., 1982). It was suggested that age-related increases in memory span arise from improvements in processing efficiency that release additional resources to support storage.

1.7.2. Time-based forgetting

An alternative view advanced by Towse and colleagues (e.g., Towse & Hitch, 1995; Towse, Hitch & Hutton, 1998) is that storage items are vulnerable to time-based forgetting while the participant is engaged in the processing requirements of the task. They argued that the Case et al. (1982) findings may have resulted from uncontrolled differences in the temporal duration of the complex memory spans rather than from trade-offs between processing and storage. They proposed that children do not simultaneously process and store materials in the course of complex span tasks, but instead switch between the processing elements of the tasks and item retention. Accordingly, the longer the processing phase of the span task, the longer the participant is switched out of remembering, and hence the more difficult it is to accurately recall memory items. Evidence consistent with this task-switching model was provided in a series of studies that either varied counting complexity while holding the overall processing difficulty constant (Towse & Hitch, 1995) or manipulated retention requirements in counting, operation and reading span tasks while holding the overall processing difficulty constant (Towse et al., 1998). The results from these experiments suggested that working memory span, rather than being a measure of capacity for resource-sharing, is constrained by a time-based loss of activation of memory items (Hitch et al., 2001).

1.7.3. Cognitive load

The multiple-factor account of working memory performance advanced by Barrouillet and colleagues (Barrouillet & Camos, 2001; Barrouillet, Bernadin, & Camos, 2004) combines concepts of both temporal decay and processing demands in a single metric of cognitive cost that is strongly related to performance on complex span tasks. The cognitive cost of a working memory span task is measured as the proportion of time over which limited capacity attentional resources are captured, for example to support memory retrievals. When attention is diverted from item storage to processing in this way, memory representations cannot be refreshed and therefore decay with time. Memory retrievals are subject to a discrete processing bottleneck that prevents simultaneous retrievals, and processing can occupy the retrieval process required to refresh the memory items. Heaviest cognitive costs and therefore lowest levels of complex span performance are therefore expected under conditions in which there is the greatest ratio of number of retrievals to units of time.

Barrouillet and Camos (2001; Experiment 3) report findings that children's complex span was higher for a task that involved articulatory suppression than one that involved mental arithmetic. The researchers argued that a complex processing task such as mental arithmetic demands sustained attention due to multiple memory retrievals whereas a 'time-filler' such as articulatory suppression does not, and therefore has a far more disruptive effect on the concurrent maintenance of items in memory due to greater temporal decay. According to Barrouillet and Camos (2001), this suggests that resource-sharing *does* occur, if only when the processing element of the span task involves increased attentional demands (such as mental arithmetic) (see also, Barrouillet et al., 2004).

1.7.4. Intrinsic storage demands

Towse, Hitch and Hutton (2002) advanced an alternative account of the Barrouillet and Camos' (2001) findings. They pointed out that mental arithmetic and articulatory suppression differ not only in the extent to which they demand effortful processing and hence limit attentional resources, but also in the extent to which they impose substantial storage demands. Whereas storage of interim products of calculations is a key feature of arithmetic calculations involving carrying operations, articulatory suppression has no obvious storage requirement. It remains an open question, therefore, whether the lower complex memory spans associated with mental arithmetic than suppression arise from storage-related interference processes, and whether resource-sharing does indeed occur under certain circumstances in working memory performance. Chapter 2 provides an investigation of the nature of the processing activity on children's complex span performance in order to examine further the potential impact of the nature of processing on recall.

1.7.5. Processing speed

The studies cited above suggest a role for both time-based forgetting and cognitive load; however, other studies have identified a range of further factors that may be important in accounting for children's working memory performance. Fry and Hale (1996) have suggested that age-related increases in processing speed underlie most of the developmental increases in working memory capacity, which in turn are a direct determinant of individual differences in fluid intelligence. Although the emphasis in this account lies on the relationship between processing speed and fluid intelligence, it is not incompatible with a task-switching hypothesis (e.g., Towse & Hitch, 1995), as faster processing would presumably enable participants to switch more rapidly from processing to storage activities, decreasing the time during which memory trace decay can occur.

Overall, it appears that complex span performance in children is mediated by a constellation of factors that undergo both quantitative and qualitative change across development. Clearly, then, there remains scope for further investigation of what mediates the relationship between processing and storage activities in children as well as adults. Furthermore, few studies appear to have investigated systematically the effects of the *nature* of processing and storage stimuli in explaining children's working memory performance. The following section reviews the literature with regard to the role of stimulus features in complex span tasks, specifically relating to the effects of similarity of processing and storage information.

1.8. Stimulus-similarity effects in working memory

1.8.1. Similarity in simple span

Evidence from studies using simple span measures indicates that span scores diminish as the similarity among items increases. For example, in a span task in which sequences of letters are presented auditorily, span performance is lower when the letters in a sequence are phonologically similar (e.g. *P G T*) than when they are phonologically dissimilar (e.g. *B Y K*; Conrad & Hull, 1964; Baddeley, 1966). In addition, span scores decrease when recall items within a sequence are from the same category (e.g. digits) compared to when the items are drawn from two distinct categories (e.g. digits and words; Young & Supa, 1941). In more recent dual-task studies, combining memory with concurrent activities that involve distinct domains, such as remembering auditorily presented words while performing an unrelated visuo-motor activity, causes only minimal task decrements (e.g., Baddeley, Bressi, Della Sala, Logie, & Spinnler, 1991; Cocchini et al., 2002).

1.8.2. Similarity in verbal and visuo-spatial span

Several studies have provided data suggesting that one of the factors mediating span performance in adults is the similarity between storage and processing stimuli (e.g. Turner & Engle, 1989; Shah & Miyake, 1996; Li, 1999), in that recall performance is worse when stimuli are taken from the same informational category (e.g., verbal processing/ verbal storage), than when stimulus categories are distinct (e.g., verbal processing/ spatial storage).

However, data regarding children's performance under experimental manipulations of stimulus similarity are rather more sparse. A notable exception is provided by Bayliss and colleagues (Bayliss et al., 2003, Experiment 1; Bayliss et al., 2005), who recently provided evidence that processing, verbal storage and visuo-spatial storage play separate roles in working memory span tasks. Specifically, Bayliss et al. showed a reduction in recall performance in the verbal domain when processing and storage were drawn from distinct domains (verbal and spatial). In contrast, Towse et al. (2002) found no such effect in children when crossing verbal processing items (nonwords and pseudohomophones) with verbal and numerical memoranda, although the authors concede that an effect may well have been found with a larger sample size.

It would seem, therefore, that a high degree of relatedness between material to be processed and stored in complex span tasks impairs complex memory span, in both children and adults. This finding is compatible with the multiple component model of working memory (Baddeley, 1986), such that when a processing task draws on visuo-spatial resources (i.e., the visuo-spatial sketchpad), recall of memory items is facilitated if those items draw on separate phonological resources (the phonological loop). In other words, it is more beneficial to task performance when concurrent tasks rely on separate systems of working memory. A stimulus-similarity effect within the verbal domain, however, is not readily explained by the multiple component model, as both processing and storage activities rely heavily on the phonological loop. In such a case, one would not expect a beneficial effect of stimulus similarity. Chapter

3 therefore provides a set of experiments investigating further the issue of stimulus similarity in children's working memory performance.

1.9. Interference in working memory

Given the robust findings of performance decrements when processing and storage stimuli are similar, the question remains as to the mechanism underlying this effect. This section addresses the question of the possible role of interference in accounting for this stimulus-similarity effect.

1.9.1. Retroactive interference

The studies described above showing evidence of stimulus-similarity effects can be interpreted in terms of a feature of working memory that has recently stimulated further theoretical development: that of interference between memory and processing items in complex span tasks. The idea that forgetting can occur as a result of interference is not a new one: McGeoch and MacDonald (1931) provided evidence for differences in serial recall of learned adjectives following the interpolation of to-be-learned material that was either similar (synonyms), or dissimilar (3-digit numbers). Adjective recall was significantly impaired when the interpolated material comprised synonyms as opposed to digits. This effect, since termed *retroactive interference*, is characterised by the number of interpolated trials with other material and the degree to which initial learning affects recall. Retroactive interference is greatest when the stimuli in the two learning tasks are the same, but the required responses are different (for a review see Anderson & Neely, 1996).

1.9.2. Proactive interference

Another interference effect that has been the focus of much research is that of *proactive interference* (PI). As with retroactive interference, PI is sensitive to the similarity of material to be processed; however, PI occurs when *previous* learning disrupts current cognitive performance. Wickens, Born and Allen (1963) were among the first to explore this topic, demonstrating that items differing from previous semantically related items were best recalled, thus yielding a release from PI. More recent research has found release from PI to be a robust phenomenon, with some researchers showing that inducing a contextual change between test trials reduces PI (e.g., Wickens & Cammarata, 1986), or that susceptibility to PI increases with age, with older adults benefiting more from reduced-PI conditions than younger adults (Lustig, May, & Hasher, 2001).

1.9.3. Interference and complex span tasks

To what extent, then, can a notion of interference explain similarity effects in complex working memory? As outlined earlier, there is abundant evidence to suggest that complex span tasks are better predictors than simple spans of performance on complex cognitive tasks such as language comprehension (e.g., Daneman & Merikle, 1996), and that complex spans capture systematic variance not contained in simple spans (Engle et al., 1999). It is possible that the unique variance associated with complex spans reflect individual differences in the ability to resist interference. Consider the task structure in a typical complex span task, the reading span task, in which a sentence is judged for plausibility, and the sentence-final word is retained for subsequent recall (e.g., Daneman & Carptenter, 1980). Participants are presented with a sentence, for example, "*Mammals are vertebrates that give birth to live young*". Every word must be read and processed, the sentence must be semantically encoded, and its veracity judged. Subsequently, the word "*young*" ceases to be a processing item and must be put aside for later recall. The remaining words "*mammals – are – vertebrates – that – give – birth – to – live*" now become obsolete. During the course of the span task, various phonological and semantic representations are generated. Those representations from the processing component of the task must then be suppressed, or inhibited, to enable accurate recall of the target memory item. In addition, in a typical span task experiment, testing commences with sets of two sentences to be verified, increasing systematically by one sentence until participants are no longer able to accurately recall the sentence-final word. Thus, participants may be presented with a total of 15 to 20 different sentences, and a corresponding number of different representations, in the course of a span task.

1.9.4. Response competition

Interference-related performance decrements are readily accounted for by an explanation in terms of PI (e.g. Lustig et al., 2001). According to such an account, the failure to suppress the obsolete items adequately increases the likelihood of competition among candidate responses, which in turn, leads to a build-up of PI. Evidence for this account was obtained recently by May, Hasher, and Kane (1999, Experiment 1), who compared older and younger adults' performance on the standard span paradigm with performance on span tasks designed to reduce the impact of interference. In the standard condition, five sets of two sentences were presented, followed by five sets of three sentences, and so on. In the reduced-interference condition, the same stimuli were used, but were presented in descending format; that is, five sets of four sentences were presented first, followed by five sets of three sentences, and so forth. Thus, the extent to which the opportunity for a build-up of PI existed was varied across the two conditions. May et al. found that older adults performed reliably better in the descending format than in the standard format, although there was no difference in the younger adults' performance across conditions. The authors argued that the findings reflect age-related impairments in the suppression of no-longer-relevant material, leaving older adults with more irrelevant information in working memory, which in turn interferes with the relevant target items. This account fits well with the stimulus-similarity effect described above, as drawing the stimuli to be processed and remembered from different representational domains (e.g., verbal and spatial) would therefore result in a decrease of PI within the span task.

1.9.5. Interference and working memory capacity

May and colleagues (e.g., May et al., 1999; Lustig et al., 2001) claimed that an interpretation of working memory function within a framework of PI is inconsistent with a capacity account (e.g., Just & Carpenter, 1992; Engle et al., 1992). Indeed, Lustig et al. (2001) state that "*Working memory span tasks may also measure ... the capacity to simultaneously store and process currently relevant information, but if so, this ability may be obscured by the presence of interference in the task*" (p. 200). It is argued here – in contrast to the May et al. interpretation – that the concept of interference can be reconciled with a notion of a limited-capacity working memory, in that working memory capacity may become further constrained when the potential for interference exists. Thus, resisting interference during a complex span task may well utilise executive resources that are required to execute other task requirements (e.g., switching between processing and storage activities, coordinating tasks, or monitoring output), to the extent that task performance suffers. It is noteworthy that when Lustig et al. (2001) refer to working memory capacity, they are referring only to the combined operations of processing and storage, and not to the additional executive functions just mentioned that may be required to perform the task successfully.

1.9.6. Executive resources

Other researchers have also speculated on a possible role of interference in executive processing. Oberauer et al. (2003) conducted a latent variable analysis to assess the distinctiveness of executive functions in working memory. They found three separable elements, of which a supervisory function was hypothesised to reflect the capacity to avoid interference from stimuli that, though previously activated, have become irrelevant during the course of the task. In a similar vein, Miyake et al. (2000) propose an updating function in executive processing, which may be responsible for inhibiting irrelevant incoming information and also deactivating or suppressing no longer relevant information.

1.9.7. Inefficient suppression mechanisms

Evidence that inefficient suppression in working memory applies not only to ageing populations and younger adults, but also to children comes from studies conducted by De Beni and colleagues (e.g., De Beni, Palladino, Pazzaglia, & Cornoldi, 1998; De Beni & Palladino, 2000; Palladino, Cornoldi, De Beni, & Pazzaglia, 2001). For example, De Beni and Palladino (2000) compared two groups of children who differed in terms of reading comprehension ability on a working memory task. De Beni and Palladino found that the groups differed in reading span performance, with poor comprehenders recalling less accurately in longer sequences of sentences. In addition, the researchers assessed the number of intrusion errors made by both groups; that is, the number of nonfinal words that were incorrectly recalled, but which belonged to the same set of sentences. Poor comprehenders produced a significantly higher number of intrusion errors, suggesting that reading span is related to suppression mechanisms in working memory.

1.9.8. The prefrontal cortex

Research in the brain sciences suggests that the frontal lobes play a significant role in the ability to effectively inhibit or suppress interference from stimuli and association that are not relevant to the task in hand (Fuster, 1997), and fMRI studies have shown that activity occurs in the dorsolateral prefrontal cortex during tasks that require the use of executive processes, such as planning, focusing attention, and task-switching (e.g., Carter, Mintun, & Cohen, 1995). A fractionation of verbal and spatial processes in the brains of nonhuman primates has been shown by neuroimaging studies (e.g., Rao, Rainer, & Miller, 1997), providing tentative support for a multi-component view of working memory, although caution must be adopted when making inferences about human behaviour.

1.9.9. Attentional control

According to the view advanced by Engle and colleagues (e.g. Engle et al., 1999), complex span tasks require controlled attention to prevent distracting secondary information from interfering with the maintenance of target memory items. In complex span tasks, some of the limited attentional resources are diverted from the memory task by the representations inevitably generated during processing. However, neither this account, nor the PI account advanced by May and colleagues, elucidates the specific mechanisms underpinning interference, nor does it explain why the greatest disruptions in memory performance arise when the processing and storage stimuli are drawn from common representational domains (e.g., Shah & Miyake, 1996; Bayliss et al., 2003).

1.9.10. Feature overwriting

A more detailed account advanced by Oberauer and colleagues (Oberauer & Kliegl, 2001; Oberauer, Lange, & Engle, 2004; Lange & Oberauer, 2005) is that interference results from partial overwriting of overlapping representations. If several distributed representations are held in working memory simultaneously, they can overwrite each other to the extent that they share some of their features (cf. Nairne, 1990). This account differs from the Engle et al. (1999) view in terms of the role of similarity of stimuli. The Engle et al. (1999)

view contends that interference occurs through distraction of attention, regardless of the similarity between representations. In contrast, in the feature overwriting account, the degree of interference is determined by the degree of overlap (i.e., similarity) between the representations of target (to be remembered) and non-target (processing) items. Saito and Miyake (2004) advanced a similar model in which stimuli generate a variety of representations (e.g., phonological, semantic, visual), and interference arises as a consequence of high degrees of featural similarity within representational domains.

However, it is important to note that the evidence regarding similarity-based interference is not unambiguous. In a study in which the similarity within both the spatial and verbal domains was manipulated, Oberauer et al. (2004) did not find consistent performance decrements in high-similarity conditions. Oberauer et al. concede, however, that similarity-based interference effects may depend on how similarity is operationalised in working memory tasks, and that there is a "*... need for theories of interference to indicate more precisely under which conditions similarity affects working memory performance.*" (p. 92). It is clear that while interference appears to play a role in complex span performance, the precise conditions under which such interference occurs remain unspecified.

1.10. Lexicality and working memory

1.10.1. Verbal serial recall

There is accumulating evidence suggesting that experimental effects that influence performance on serial recall tasks are also present in complex memory paradigms. For example, LaPointe and Engle (1990) reported that the length of the words to be recalled in a complex span task negatively influenced recall. The finding of a word length effect in such a task suggests that verbal complex span is sensitive to a manipulation shown to influence recall within simple span tasks. In a more recent study, Lobley, Gathercole and Baddeley (in press) provided evidence of the ability of phonological similarity – an effect known to impair serial recall – to reduce recall performance on a listening span task. It would therefore appear that serial recall and complex memory span paradigms tap some common cognitive processes.

1.10.2. Words and nonwords

One way of investigating whether working memory and short-term memory are supported by common cognitive processes, and to examine in more detail the notion of similarity-based interference in complex span tasks, is to make use of a further effect observed in serial recall tasks. One such effect is the so-called lexicality effect: the recall superiority for lists of words over nonwords (e.g., Gathercole, Pickering, Hall, & Peaker, 2001; Hulme, Maughan, & Brown, 1991). There is evidence to suggest that short-term retention of verbal material is mediated not only by a memory system dedicated to holding verbal information for brief periods (such as Baddeley's (1986) phonological loop), but also that recall of such information is supported by LTM. That is, words are highly practised, familiar stimuli that have phonological and semantic LTM representations. In contrast, nonwords are unfamiliar and presumably lack corresponding representations in LTM.

1.10.3. Lexicality and interference

How does lexicality relate to interference in complex span tasks? With regard to a possible role for feature overwriting in complex span tasks under conditions of stimulus similarity, words and nonwords may provide a useful stimulus set with which to examine more closely the disruptive consequences of similarity between processing and storage items. Whereas words and nonwords both generate phonological representations during the course of a span task, words differ from nonwords in that they generate additional lexical-semantic representations. Thus far, the lexicality of target memory items does not appear to have been studied in complex memory span. Hence, it remains an open question whether similarity-based interference within the verbal domain operates at a lexical-semantic as well as a phonological level. In addition, given the robust effect of lexicality in healthy participants, it is possible that lexicality exerts a beneficial influence on complex span performance, which would provide additional support for evidence that serial recall and complex memory span paradigms tap some common cognitive processes (e.g., LaPointe & Engle, 1990; Lépine, Barrouillet, & Camos, 2005; Lobley et al., in press), especially given recent evidence that measures of semantic short-term memory are a more reliable predictor of comprehension than traditional phonological short-term memory measures such as simple word span (Haarman et al., 2003).

1.11. Summary and Aims of the present study

In summary, the review presented here has highlighted the diversity of approaches in investigating human short-term memory. Various theoretical working memory models have been described and contrasted, and their empirical support has been evaluated. With regard to measures of working memory, it appears that the complex span paradigm remains a highly popular tool for assessing working memory performance; its popularity due both to its reliability and its predictive power in connection with performance on 'real-world' cognitive tasks.

While the majority of researchers would agree that complex span appears to be a complex phenomenon drawing on many levels of representation, there remains little consensus over the detailed nature of the relationship between processing and storage activities in measures of working memory. Whereas some researchers argue that these processes are dynamically coupled in the sense that they compete for a single, flexible resource, others suggest that processing and storage are not in direct competition, but are nonetheless linked, possibly in terms of temporal resources. The concept of interference has been linked to executive processing, but there are inconsistencies as to when interference between processing and storage stimuli actually occurs. Finally, the review highlights that a potential influence of lexicality on complex span performance has not, thus far, been adequately addressed.

The primary aim of this thesis is to further examine the nature of the relationship between processing and storage activities in children's and adults' working memory. Chapter 2 presents a set of three experiments that examine the impact of the processing task on children's recall performance in complex span tasks. Experiments 4 and 5 (Chapter 3) provide a systematic investigation of the similarity of processing and storage stimuli on children's working memory performance. The final experimental chapter (Experiments 6-8) focuses on the effects of the lexical status of processing and recall items in both adults' and children's working memory tasks. The final part (Chapter 5) summarises the findings from the eight experiments and provides a critical discussion of the data in the light of existing theory. The thesis concludes with a consideration of the overall findings with regard to their theoretical implications.

Chapter 2

Processing demands in children's span task performance

2.1. Introduction

One influential account developed by Case et al. (1982) to explain developmental differences in working memory performance across the childhood years is that working memory is a single flexible system fuelled by a limited capacity resource that can be flexibly allocated to support processing and storage. By this view, the total working memory resource remains constant, but the efficiency of processing speed increases with age. In Case et al.'s (1982) study, storage space was measured independently of processing efficiency by using a counting span procedure: participants were required to count target objects on a series of cards and then recall all counted totals. A participant's counting span was the maximum set size for which he or she could recall all the count totals, on at least two out of three trials. Case et al. showed that with children between the ages of six and twelve years, there was a positive correlation between counting span and counting speed. In order to show that increases in counting efficiency were responsible for increases in short-term span, and that counting speed could predict counting span. Case et al. went on to manipulate task difficulty with adult participants, by using a counting span paradigm in which the participants were required to count using a set of learned nonsense numbers (e.g. *rab, slif, dak*). The subsequent measure of adults' counting speed and span showed that under these conditions, adults' spans were not significantly higher than those of the tested six-year olds. In all experiments, the results showed a linear relationship between span and speed, seemingly supporting Case et al.'s view that operating and storage functions compete for common resources within a total processing space. According to this "trade-off" hypothesis, the younger children were less efficient at counting, thereby utilising more operating space for the counting procedure and leaving less space for storage. When adults' counting efficiency was reduced by making them count using nonsense words, they also experienced a reduction in storage space. Hence, the effect of the complexity of the processing task is explained in terms of resource-sharing within a unitary working memory capacity.

Towse and colleagues (Towse & Hitch, 1995; Towse et al., 1998; Hitch et al., 2001; Towse et al., 2002) have criticised the resource-sharing interpretation, arguing that Case et al.'s (1982) findings may have resulted from uncontrolled differences in the temporal duration of the complex memory span tasks rather than trade-offs between processing and storage. In Case et al.'s (1982) study, increases in counting difficulty involved a corresponding increase in the time during which card totals had to be held in memory. In other words, Case et al. equated processing speed with processing efficiency. Towse and colleagues argued that the improved counting spans of older children was due to their counting more quickly than young children, and proposed that children do not simultaneously process and store material in the course of complex span tasks, but instead switch between the processing element of the task and item retention. By this account, poorer span performance under more complex processing conditions results from the greater opportunity for time-based forgetting due to the lengthier retention intervals.

Evidence consistent with this task-switching model was initially provided by Towse and Hitch (1995) in a study that, similar to Case et al. (1982), attempted to measure counting span in children between the ages of five and eleven years. However, in Towse and Hitch's study, counting span was measured in three different conditions. In the first condition – the "feature" condition – children were required to count target objects that were easily distinguishable, on the basis of their colour, from non-target objects in an array. In the second – "conjunction" – condition, the target objects were not easily distinguishable from the non-target objects (the colour of both types of object was the same). The third type of array was called the "feature-slow" condition, and was constructed in the same way as the feature cards (i.e. target objects possessed a unique colour), the difference being that feature-slow arrays contained a greater number of target objects. The logic behind the different arrays was to vary task difficulty between conjunction and feature-slow arrays, while holding processing time (counting time) constant. As predicted by Towse and Hitch, the results showed that children's counting spans differed significantly between feature-slow and conjunction arrays. Towse and Hitch argued that these results demand an alternative explanation to Case's (1985) resource-sharing hypothesis, as this would have predicted a difference in span between the feature-slow and conjunction conditions on the basis of differences in processing requirements.

In a further series of studies, Towse and colleagues went on to manipulate retention requirements in counting, operation and reading span tasks, while holding constant the overall processing difficulty (Towse et al., 1998). The results from these experiments suggested that rather than being a measure of capacity for resource-sharing, working memory span is constrained by a time-based loss of activation of memory items (Hitch et al., 2001; Ransdell & Hecht, 2003).

Further evidence for the notion that children's working memory comprises separate systems for processing and storing material, rather than a single flexible capacity that deals with both processing and storage demands, comes from a study by Halford, Maybery, O'Hare, and Grant (1994). In a modification of the counting span task, Halford et al. (Experiment 3) requested their five- to twelve-year old participants to remember a pre-load before counting target objects on sets of cards, and then to recall the pre-load (rather than the count totals). By holding the pre-load (i.e. the storage requirement) constant, the researchers hoped to tease out any potential trade-off effects that might occur between storage and processing operations. The results showed that memory declined as a function of the number of cards featuring count arrays, and have two possible interpretations: if counting and pre-load storage had to compete for resources from the same processing capacity, as would be predicted by a resource-sharing model, then the younger children should either show a greater decline in pre-load recall (that is, they count at the expense of storing), or a decline in counting rate (that is, they store the pre-load at the expense of counting efficiency). However, younger children failed to show a significant decline in pre-load recall as a function of cards counted, despite the fact that they showed no reduction in counting rates between pre-load and no-pre-load conditions. These findings undermine Case's (1985) resource-sharing hypothesis, suggesting instead that storage and processing activities in working memory are supported by two distinct systems.

Three experiments were designed to investigate further the extent to which the nature of processing influences children's performance on complex memory span tasks. The experiments also investigate whether the impact of processing activities on memory span is subject to developmental change. The two age groups were included in order to test for the generality of the experimental findings across age, in line with other studies that have examined working memory performance in children (e.g. Halford et al., 1993; Towse & Hitch, 1995; Barrouillet & Camos, 2001). Findings of age-related changes in the factors influencing complex memory span would provide new insights into the nature of developmental changes in working memory function, the nature of which is not at present fully understood.

2.2. Experiment 1

The first experiment extends the approach adopted by Towse and Hitch (1995) to investigate the influence of processing complexity on performance in a complex memory span paradigm involving mental arithmetic. Children aged seven and nine years participated in the study, and were required on each trial to add sequences of numbers and recall each total for later serial ordered recall. The calculation varied in difficulty across two conditions, involving the addition of either of single digit numbers or 2-digit numbers that required a carrying operation. The time taken to complete the calculations in the two tasks was equated by presenting longer sequences of numbers for addition in the single- than double-digit conditions.

Resource-sharing and temporal decay accounts of working memory make contrasting predictions concerning the outcomes of this experiment. According to a resource-sharing account, memory span should be lowest for the calculations involving carry operations, as the resources available to support item retention will be diminished in this condition as a consequence of the greater processing load. In contrast, by a temporal decay account, memory spans should be equivalent for conditions involving carry and simple operations, as their temporal durations are equivalent.

2.2.1. Method

2.2.1.1. Participants

A total of 64 children from Year 3 (*N* = 33, mean age 7 years 9 months, range 7;4 to 8;3) and Year 5 (*N* = 31, mean age 9 years 9 months, range 9;3 to 10;3) of a local primary school in Stockton-on-Tees, UK, participated in the experiment. Participants were taken from a sample of children who were identified by their teachers as having normal arithmetic skills.

2.2.1.2. Design

The experiment employed a two-way mixed design with age as a between-subjects factor (7 and 9 years) and type of operation as a within-subjects factor (simple and carry sums). Dependent variables were the number of operation totals recalled accurately (operation span), the time taken to calculate operations (operation speed), and calculation errors. The order of testing the two conditions was counterbalanced.

2.2.1.3. Materials

In the carry condition, problems consisted of the addition of two 2-digit numbers that involved a carry operation of the units, e.g. *35 + 17*. The simple condition involved the addition of a series of five single digit numbers, e.g. *1 + 2 + 1 + 2 + 3*. Pre-tests with both age groups (sample of 5 children from each age group) allowed these simple problems to be matched for time with the carry problems.

2.2.1.4. Procedure

All children were tested individually in a quiet area of their school. A laptop computer with a 12-inch colour monitor was programmed to control the display of individual operations and to record the response times. Totals recalled subsequently were recorded on score sheets. The children were told that they would be shown a sum on the computer screen that had to be worked out, and that as soon as they had reported the answer out loud, another sum would appear, which would also need to be calculated and reported. They would then be requested to recall, in order, the totals previously calculated. It was emphasised that although they were being timed, it was important that they try and work out the answer as accurately as possible. A practice trial at the beginning of each task established that all the children grasped the concept immediately.

Each condition began with a sum displayed on the screen, centre-justified, as black numbers on a white background in 72-point Arial font. As soon as the answer had been reported, the next operation appeared, initiated by a key press by the experimenter. After each calculation, the answers were recorded manually and the response times were recorded electronically on the laptop following a key press by the experimenter. The participant was then asked to recall, in order, the successive totals. These were also recorded. Correct responses were scored in terms of the total that had been calculated, not the actual total. So, if a child erroneously gave the answer "*50*" to the operation "*35 + 17*", and then subsequently recalled "*50*", a correct response was recorded. If a child was successful in recalling totals from two trials (out of a possible three), the number of operations to be calculated—and therefore totals to recall—was increased by one. However, if on more than one out of three trials the child did not recall the totals correctly, the span testing was discontinued for that condition. After a short break, testing resumed in the other condition. Operation span (in this and subsequent experiments) was calculated as the maximum level at which recall was correct, with 0.5 points added if a single trial at the next level was also correct. In addition, the number of correct answers (correct item in the correct serial position) in each remaining trial was calculated as a proportion of the number of items to be recalled. This value was multiplied by 0.5 and the product added to the total score obtained from the procedure above.

2.2.2. Results

The data for three children from the younger age group who were unable to calculate the carry operations was excluded from the analysis, leaving scores for 31 participants in the 7-year old age group. Mean memory spans for the two types of operation, as well as reaction times and error rates, are shown in Table 1.

TABLE 1

Mean span performance, reaction times (in seconds), and calculation errors (per 100 operations) and standard deviations for the 7- and 9-year-olds for simple and carry sums in Experiment 1

	Simple sums				*Carry sums*			
	7 years		*9 years*		*7 years*		*9 years*	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Span	1.69	0.41	2.20	0.61	1.63	0.41	2.15	0.71
RT	14.57	3.91	7.80	1.97	14.68	3.88	7.72	2.62
Errors	7.89	12.33	4.16	6.07	22.53	22.06	18.99	20.77

Memory spans were higher for the 9- than 7-year old groups, but did not vary markedly as a function of processing operation. A two-way analysis of variance as a function of age and operation type was performed on the span scores. The analysis yielded significant main effect of age, $F(1,59) = 22.48$, $MSe = 8.16$, $p < .001$, partial $\eta^2 = 0.276$, with mean operation span higher in the 9-year-old children (2.18) than in the 7-year-olds (1.66), but no significant effect of span, $F(1,59) = 0.37$, $MSE = 0.09$, $p > .05$, partial $\eta^2 = 0.006$, and no significant interaction, $F(1,59) < 1$, partial $\eta^2 = 0.001$.

A corresponding analysis of variance was performed on the calculation speed scored for each child as a function of duration and age. There was no significant difference in speed as a function of type of operation, $F(1,59) = 0.002$, $MSe = 0.06$, $p > .05$, partial $\eta^2 = 0.001$. The older children were significantly faster than the younger children on this measure, $F(1,59) = 929.62$, $MSE = 15287.73$, $p < .001$, partial $\eta^2 = 0.94$. The interaction between operation type and age was nonsignificant, $F < 1$, partial $\eta^2 = 0.001$.

Error rates were higher in the carry condition (20.67 %) than in the simple condition (6.03 %). Due to the non-normal distribution of error rates, these data were tested statistically using the Wilcoxon test. The increased rates of error in the carry than simple sum conditions were statistically significant for both the 7-year olds, $z = 4.38$, $p < .001$, $r = 0.56$, and the 9-year olds, $z = 4.10$, $p < .001$, $r = 0.52$. Finally, the correlation between processing speed and mean span for both types of task was significant, $r = -.43$, $p < .001$. This finding replicates previous findings (e.g. Hitch et al., 2001; Towse et al., 1998) that processing times were related to storage, and is consistent with both a trade-off and time-based forgetting accounts.

2.2.3. Discussion

More errors were made for the calculations involving carrying sums than simple sums, indicating that the computations were more difficult. However, span did not differ across the two conditions, and calculation speed was equivalent between the two conditions. These findings are not consistent with a simple resource-sharing account of working memory, according to which operation span should decrease in the carry condition as a consequence of the increased processing demands of the task. The results from this experiment favour instead the task-switching account advanced by Towse and colleagues (e.g. Towse & Hitch, 1995; Hitch et al., 2001). According to this, in the course of complex memory span tasks individuals alternate between processing (in this case, performing an arithmetic calculation) and storage. An important factor in memory performance is the time taken to carry out the processing activity, during which the memory representations are lost (either through decay or some other forgetting mechanism). As processing time was equivalent for both types of operation in the present experiment, comparable levels of performance in the two conditions would indeed be expected.

Findings by Barrouillet and Camos (2001; Experiment 3) do, however, suggest that processing demands play a role in complex span performance. They argued that the manipulation of counting duration in the Towse and Hitch (1995) study may have led to changes in the cognitive cost of the task. One assumption underlying the Towse and Hitch study was that the difference between counting large and small arrays is the time taken to complete the task. However, it is possible that the counting of larger numbers of items constitutes a more demanding task. Developmental research suggests that both pointing and counting performance is greatly influenced by variations in the number of objects (Gelman & Meck, 1983; Potter & Levy, 1986; Camos, Fayol, & Barrouillet, 1999). Thus, the researchers argue, the Towse and Hitch findings may be interpreted in terms of differences in cognitive demands between counting conditions.

In order to test this possibility, Barrouillet and Camos (2001; Experiment 3) conducted a study in which the duration of the processing activity was held constant, but in which the cognitive cost of the task was manipulated. They compared children's performance on an operation span task with a task in which children were merely required to suppress articulation for a corresponding period of time. Complex span for consonants was significantly poorer when the intervening activity involved arithmetic calculations than when it involved articulatory suppression. On this basis, Barrouillet & Camos suggested that a critical factor constraining performance on complex span tasks is the extent to which the processing task is demanding of attention over a set duration (see also, Barrouillet et al., 2004; Gavens & Barrouillet, 2004). Specifically, they argued that mental arithmetic demands sustained attention due to multiple memory retrievals whereas articulatory suppression does not, and therefore has a far more disruptive effect on the concurrent maintenance of items in memory due to greater temporal decay. According to Barrouillet & Camos (2001), this suggests that resource-sharing *does* occur, if only when the processing element of the span task involves increased attentional demands (such as mental arithmetic). In their interpretation, the automatized nature of counting does not require more attentional resources than suppressing articulation, resulting in similar spans for these two conditions.

However, in their experiment, the processing activities in the two conditions differed not only in terms of attentional demands, but also in terms of intrinsic storage demands. As noted by Towse et al. (2002), articulatory suppression does not require the retention of interim solutions in the way that mental arithmetic does. It is possible that the lower complex memory spans associated with mental arithmetic than suppression arise from storage-related interference processes.

2.3. Experiment 2

A second experiment was designed to distinguish between the influences of processing and intrinsic storage demands of interpolated tasks on complex memory span performance in children. Memory span was compared for three processing tasks that varied in their processing and intrinsic storage demands: Mental arithmetic involving carry operations imposes significant demands on both attention and storage. Articulatory suppression involving the repeated production of a single verbal item requires minimal processing and no intrinsic storage. A third processing activity imposes significant demands on attention but not on storage; this task involves judging whether each of a series of 2-digit numbers are odd or even, and requires access to stored knowledge of the numerical status of each digit, but no short-term storage of successive numbers or of their odd/even status.

According to resource-sharing accounts such as Case (1985), memory span is inversely related to processing difficulty and therefore should be lowest in the mental arithmetic condition, higher in the odd/even condition, and greatest in the articulatory suppression condition. In contrast, the attentional resources account (Barrouillet & Camos, 2001) would predict equivalent performance in the odd/even and operation span conditions, as the attentional demands of these two tasks are both significant. The suppression condition, however, should yield the highest span, as there are little or no processing requirements in this condition. In contrast, a finding that performance was greater in the odd/even condition than the mental arithmetic condition, would be entirely consistent with the proposal by Towse et al. (2002) that memory span is impaired under conditions in which the processing task has its own competing memory demands. In order to investigate whether differences in processing requirements, rather than differences in the time taken to execute the task, lead to differences in span, the period during which children were engaged in the processing activities was equal across all three conditions. Once again, age groups of children were tested (aged 7/8 years and 9/10 years) in order to establish the generality of findings across age.

2.3.1. Method

2.3.1.1. Participants

A total of 63 children from Year 3 (*N* = 32, mean age 7 years 7 months, range 7;5 to 8;4) and Year 5 (*N* = 31, mean age 9 years 7 months, range 9;3 to 10;1) of a local primary school in Stockton on Tees, UK, participated in the experiment. The children had not taken part in Experiment 1.

2.3.1.2. Design

A two-way mixed design was employed, with age (7 vs. 9 years) as a between-subjects factor, and type of interpolated task (arithmetic, odd/even judgement, articulatory suppression) as the within-subjects variable. Span scores were calculated as in Experiment 1. The number of errors in the mental arithmetic and odd/ even conditions was also scored.

2.3.1.3. Tasks and Procedure

Following the task design in Barrouillet and Camos' (2001) study, tasks were administered in two sessions, in order to match exactly the duration of the individual processing tasks. Therefore, the task administered in session one was always the operation span task; the odd/even task and articulatory suppression task were conducted in session two, and the design was counterbalanced by task order for these two tasks. The sessions were three weeks apart. Each session lasted a maximum of twenty minutes and began with a practice task. Figure 1 illustrates the task design with the interpolated task requirements.

Mental arithmetic span task. For this task, the stimuli were taken from the carry sums in Experiment 1, that is, problems consisted of the addition of two 2-digit numbers that involved a carry operation of the units, e.g. *28 + 16.* A sum was displayed on a computer screen, which the child was requested to work out. As soon as the answer had been given out loud, a new problem appeared on the screen (following a key press from the experimenter). Reaction times for each operation were recorded electronically. At the end of a series, the child was requested to recall, in order, the answers calculated.

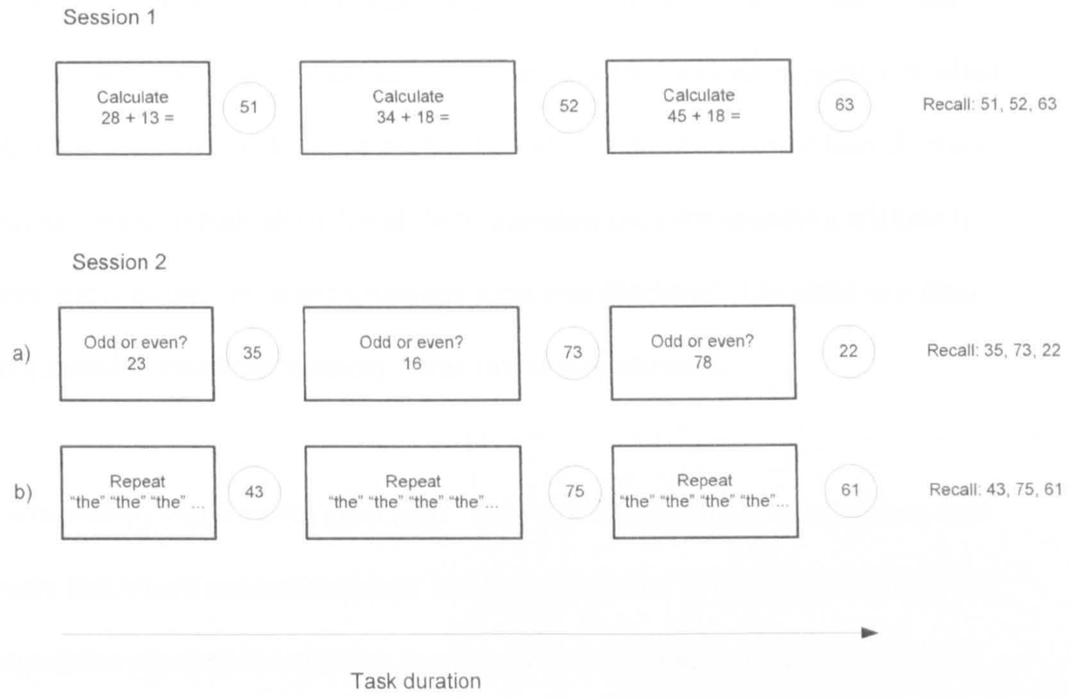

Figure 1. Schematic representation of the task design in Experiment 2. The rectangles represent the interpolated task: Session 1 requires calculation of addition operations; Session 2a) requires the participant to decide if the number presented is odd or even; Session 2b) requires the participant to suppress articulation by repeating the word "the". The circles show the items to be remembered and recalled. The interpolated tasks in Session 2 are temporally equivalent to the task in Session 1.

Odd/even span task. For the odd/even task, a series of 2-digit numbers (randomly generated by the computer) were presented on the computer screen for a period of 1 *s* each (black numbers on a white background). The child was required to state out loud whether each number was odd or even. The duration of the series was determined by the time the child had taken in the corresponding arithmetic task. For example, if it had taken a child ten seconds to calculate the sum *28 + 16* in the arithmetic span task in Session 1, then a series of numbers (judged by the child to be odd or even) would be displayed on the computer for 1 *s* per number for a total of 10 seconds. At the end of this time period, a black number was displayed prominently on a red background for 1.5 seconds. This was the item to be remembered and subsequently recalled. Another series of random numbers to be judged odd or even was then displayed for as long as it had taken that child to calculate the corresponding arithmetic operation, before the second memory item was displayed. The child was then requested to recall the memory items in the correct order.

Articulatory suppression span task. This task differed from the odd/even task only that where random numbers had been presented in the odd/even task, this condition required the children to suppress articulation by repeating the word "*the*" (at approximately one "*the*" per second) while looking at a blank screen. Again, the suppression duration was matched in time with the individual child's corresponding arithmetic calculation duration. At the end of this time period, a two-digit number was presented on the screen for 1.5 seconds, before the screen went blank again and the child was again required to suppress articulation. Thus, for each child in each series, the retention period was identical in all three tasks.

2.3.2. Results

Table 2 shows mean span performance (and error rates for the odd/even and operation span conditions) of the 7- and 9-year olds for the different types of interpolated task.

TABLE 2

Memory span performance (and standard deviations) of the 7- and 9-year olds in Experiment 2 for the different types of interpolated tasks, as well as processing errors (per 100 items) for the odd/even and mental arithmetic conditions.

Interpolated task	*7 years*		*9 years*	
Memory span	*M*	*SD*	*M*	*SD*
Mental arithmetic	1.64	0.39	2.17	0.70
Articulatory suppression	2.30	0.44	2.77	0.54
Odd/even judgement	1.32	0.25	1.72	0.34
Processing errors				
Mental arithmetic	22.27	21.69	10.97	12.15
Odd/even judgement	3.94	6.05	3.05	4.58

Memory span scores were greatest for the articulatory suppression condition, at an intermediate level for the mental arithmetic condition, and lowest in the odd/even condition, for both age groups. A two-way mixed ANOVA as a function of age and interpolated task was performed on the span scores. The results show a significant improvement in span with age, $F(1,61) = 2812.03$, $MSe = 744.27, p < .001$, partial $\eta^2 = 0.98$, and a significant main effect of task, $F(2,122) = 84.51$, $MSe = 16.30, p < .001$, partial $\eta^2 = 0.58$, but no significant interaction, $F(2,122) = 0.36$, $Mse = 0.07, p > .05$, partial $\eta^2 = 0.006$.

Simple effects of task were explored in a series of one-way within-subjects analyses of variance for each age group. The effect of task was significant for both the younger children, $F(2,62) = 65.28$, $MSe = 7.73, p < .001$, partial $\eta^2 =$ 0.68, and the older children, $F(2,60) = 31.98$, $MSe = 8.63, p < .001$, partial $\eta^2 =$ 0.52. Planned contrasts showed that for each of the age groups, span performance on the odd/even span task was significantly poorer than both the arithmetic span task and the articulatory suppression task (*ps* < .01), and that performance on the articulatory suppression task was also significantly better than the arithmetic span task (*p*s $< .01$).

Errors for the odd/even and operation conditions were analysed using the Wilcoxon test, due to the nonnormal distribution of the data. There were significantly more processing errors in the arithmetic condition (16.71 %) than in the odd/even condition (3.50 %), $z = 5.00, p < .001, r = 0.63$. This difference was significant for both the 7-year olds, $z = 3.93, p < .001, r = 0.69$, and the 9-year olds, $z = 3.03, p < .01, r = 0.54$, and indicates that the level of complexity was higher for the mental arithmetic than the odd/ even task.

2.3.3. Discussion

In both age groups, span performance varied significantly according to the nature of activity performed during the interval between memory items, despite the temporal equivalence of conditions. The articulatory suppression condition yielded higher spans than both the mental arithmetic and odd/even conditions, and lowest levels of performance were found in the odd/ even condition.

The span advantage when the interpolated task involved articulatory suppression compared with mental arithmetic replicates Barrouillet and Camos' (2001) findings, and is consistent with their view that attentionally-demanding processing activities divert limited attentional resources from storage and hence lead to accelerated temporal decay (Barrouillet et al., 2004). However, the lower levels of span performance observed in both age groups in the odd/ even than the mental arithmetic conditions do not readily fit with any existing theoretical account. First, as both processing activities are attention-demanding and mental arithmetic to an extent that is at the very least equivalent to and probably more demanding than the odd/even judgments, either comparable levels of performance or an advantage to the odd/ even task would be expected according to Barrouillet et al. (2004). Second, and relatedly, the decrement in odd/even span cannot be explained in terms of greater processing demands leading to reduced availability of storage according to a trade-off account (Case, 1985). Third, the span advantage to mental arithmetic over odd/ even cannot be explained in terms of differences in intrinsic storage demands (Towse et al., 2002), as these are greater in the former than the latter tasks. Finally, the temporal equivalence of all three processing conditions rules out any account in terms of differences in temporal decay (Towse & Hitch, 1995).

One possibility is that the unexpected finding of lower span scores in the odd/ even processing condition than in the mental arithmetic condition may have reflected differences in task structure rather than processing or storage demands. Whereas the mental arithmetic condition was self-paced, participants in the odd/even condition participants were forced to make parity judgements at an externally-determined rate of one number per second. There is recent evidence that external pacing does have a more disruptive effect on complex span than self-pacing, probably due to its disturbance of optimal switching strategies (Barrouillet et al., 2004).

2.4. Experiment 3

Experiment 3 was conducted in order to determine whether the differences in memory span across the mental arithmetic and odd/even conditions would persist if the pacing requirements of the two processing activities were equated. In Experiment 2 only the mental arithmetic condition was self-paced. In Experiment 3, both the mental arithmetic and odd/even tasks were self-paced, with presentation of successive items for processing initiated by the participant's response to the previous item. A finding that the performance cost to odd/ even judgements over mental arithmetic persists in this experiment would rule out the possibility that differences between these two conditions in Experiment 2 reflected the varying pacing requirements of the tasks.

2.2.1. Method

2.4.1.1. Participants and design

A group of 9 and 10-year old children was recruited (*N* = 42, mean age 9;8, range 9;1 to 10;2) from a local primary school to participate in the experiment. In the absence of any age-related interactions in the previous experiments, the sample comprised children of a single age group. Type of interpolated processing task (arithmetic, odd/even judgement) was the independent variable. Span scores were calculated in each condition, and additional measures taken of numbers of items processed and processing accuracy.

2.4.1.2. Materials and procedure

The experiment was conducted using a laptop computer, programmed to control presentation durations. The tasks were presented on a 13 inch colour monitor; the processing items were coloured black, and the storage/memory items were coloured red. All items were presented in black 72-point Arial font against a white background. The interpolated processing task took place within an 8-second window, with a response-based presentation format. Specifically, in the mental arithmetic condition, the child was presented with a simple operation (e.g. *12 + 3 =?*) and required to calculate the answer, with a further number presented for addition each time the total was spoken aloud (following a key press by the experimenter). This allowed for continuous processing throughout the 8-second window. At the end of the processing phase, a 2-digit number (randomly generated by the computer) was displayed prominently in red for 2 s on the screen. This was the item for retention and subsequent recall. Another processing phase then commenced for a duration of 8 *s*, followed by the presentation of a further memory item. As before, trials were increased by one if two out of three items were correctly recalled. Similarly, in the processing phase of the odd/even condition, numbers were presented in reaction to the child's spoken response, for an overall maximum of 8 *s*.

2.4.2. Results and discussion

Mean span in the mental arithmetic condition was 2.07 ($sd = 0.68$) and mean span in the odd/even condition was 2.19 ($sd = 0.72$). No significant differences in memory span were found across the two conditions, $t(41) = 1.33, p > .05$, $d = 0.17$. In addition, there were no differences in the number of items processed in either of the conditions, $t(41) = 1.71, p > .05, d = 0.17$ [mean number of additions: 5.62 ($sd = 1.55$); mean number of digits assessed: 6 ($sd = 1.38$)], indicating that the tasks captured attention in a comparable manner.

Across these two conditions, the same number of items was processed in the same period of time under conditions of self-pacing. In the odd/even task, such retrievals took the form of accessing stored knowledge of the numerical status of each digit, whereas in the operation span task each sum involved a sequence of simple additions. The elimination in the present experiment of memory span differences across these two conditions found in Experiment 2 indicates that they arose from variations in task pacing.

It should be noted that the mental arithmetic task involving carrying operations in Experiment 2 was changed to that of successive addition of simple sums in Experiment 3, corresponding to the procedure adopted by Barrouillet et al. (2004). In line with Barrouillet et al.'s own findings that the complexity of arithmetic computations *per se* has no effect on complex span, span levels were very similar for the 9/10 year old group in Experiment 2 (2.07) and Experiment 3 (2.17).

Finally, a correlational analysis of the association between span scores and number of items processed was conducted in order to examine whether the length of time taken to process individual items was linked to span performance. The two measures were highly correlated with one another, $r(40)=.70, p<.01$, indicating that children who processed most items typically had higher memory spans. This suggests that the children who took longer to process individual items did not utilise the allocated 8 *s* processing time to rehearse or otherwise consolidate memory items.

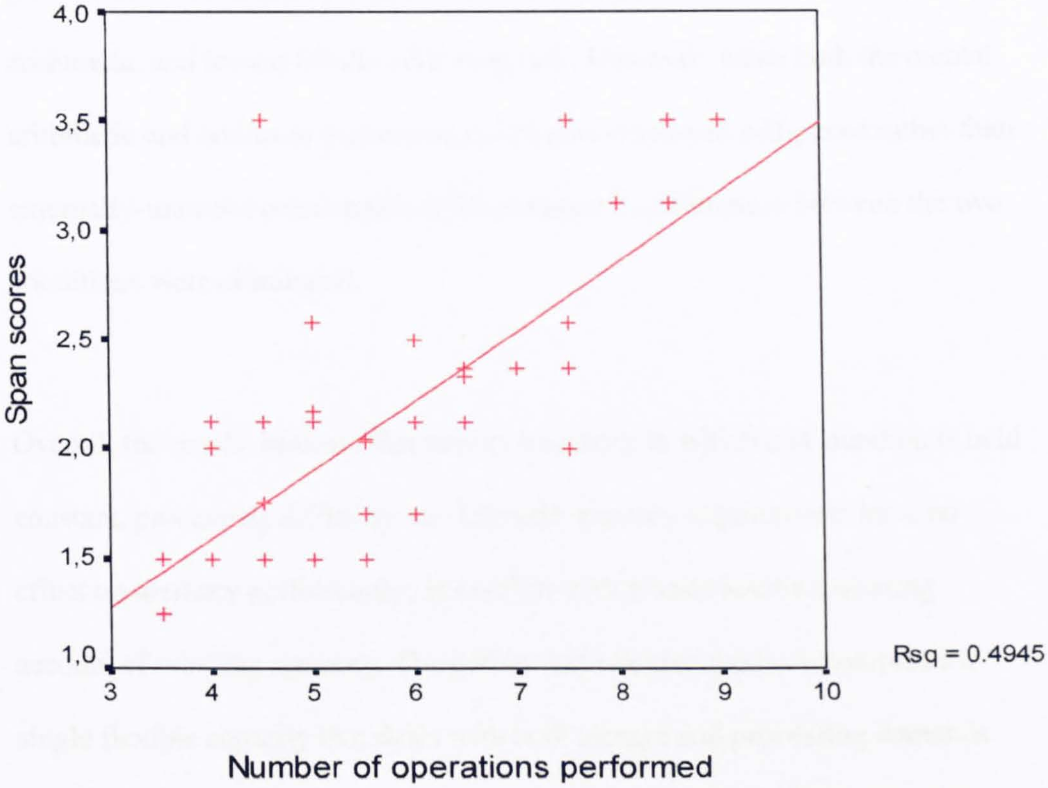

Figure 2. Individual span scores as a function of the total number of operations performed in the arithmetic and odd/ even conditions

2.5. Chapter summary

Three experiments were designed to investigate the cognitive processes involved in children's complex working memory span, by manipulating the nature of the processing activity. In all three experiments, the time spent on the processing activity prior to recall was held constant, and the complexity and intrinsic memory demands of the processing activities manipulated. In Experiment 1, span scores were found to be independent of the difficulty of mental arithmetic operations, with carry and simple sums yielding comparable spans despite differences in task difficulty as indexed by performance accuracy. In Experiment 2, three different processing activities were compared - mental arithmetic, odd/even judgements, articulatory suppression. Span scores were greatest in the articulatory suppression conditions, intermediate for mental arithmetic, and lowest for the odd/ even task. However, when both the mental arithmetic and odd/even processing conditions employed self-paced rather than externally-imposed presentation in Experiment 3, differences between the two conditions were eliminated.

Overall, the results indicate that under conditions in which task duration is held constant, processing difficulty and intrinsic memory requirements have no effect on memory performance, in conflict with a basic resource-sharing account of working memory. The notion that working memory comprises a single flexible capacity that deals with both storage and processing demands (e.g. Case, 1985) cannot accommodate the absence of a task difficulty effect in Experiment 1, thereby challenging the notion that a more difficult task will result in a greater consumption of limited cognitive resource and hence a reduction in capacity for storage.

In Experiments 1 and 3, memory span performance was equivalent across different processing conditions conducted over matched time periods. While this aspect of the results fits well with claims that storage period and hence opportunity for time-based decay are important (e.g., Towse & Hitch, 1995), it is clear that span is constrained by other factors too. Performance was greater in the articulatory suppression condition than either the mental arithmetic or odd/ even processing conditions of Experiment 2, replicating earlier findings of Barrouillet and Camos (2001), and Barrouillet et al. (2004). The nature of the processing activity therefore clearly has significant consequences for complex span, in children as in adults.

Chapter 3

Stimulus similarity decrements in children's working memory span

3.1. Introduction

The previous experiments demonstrated that the nature of processing in complex span tasks can have an effect on recall performance, indicating a link between processing and storage in working memory measures. As reviewed earlier, several studies have provided data suggesting that one of the factors mediating span performance in adults is the similarity between storage and processing stimuli (e.g. Turner & Engle, 1989; Shah & Miyake, 1996; Li, 1999).

There is, however, less evidence regarding the effects of stimulus similarity on children's working memory performance. An important exception is provided by Bayliss et al. (2003), who conducted a study with 8- and 9-year old participants (Experiment 1) and adults (Experiment 2). Four complex span tasks were developed by crossing verbal and visuo-spatial processing tasks with verbal and visuo-spatial storage requirements. In the children's task, participants were presented with a display of nine different coloured circles, each circle containing one of the digits 1 to 9. In the verbal processing task, participants were presented auditorily with a series of object names and were required to identify the colour typically associated with each object name. The visuo-spatial processing task required the children to locate within the display a visually distinctive circle. In the verbal storage condition, the children recalled the digits that were displayed in the circles; in the visuo-spatial storage condition, the locations of the target circles were recalled. Thus, children were required to either a) verbally associate colour and item prior to recalling digits, b) verbally associate colour and item prior to recalling circle locations, c) locate visually distinct circles prior to recalling digits, or d) locate visually distinct circles prior to recalling circle locations. The adult tasks in the second experiment were similar, except that the visuo-spatial processing task was modified to increase the task demands by using a conjunctive search task, in which target items had two features for identification (size of target item, and whether the item had a bevelled or unbevelled edge).

Bayliss et al. (2003) found that performance on the span tasks was dependent on the particular combination of processing and storage involved. Specifically, when verbal processing was combined with verbal storage, significant span decrements were observed, demonstrating a stimulus-similarity effect when information from within a single domain must be combined in a complex span task. However, there was no corresponding effect observed for visuo-spatial processing and storage tasks, suggesting either that visuo-spatial tasks rely exclusively on domain-general working memory resources (as opposed to verbal complex span tasks, which reflect the use of domain-specific storage and domain-general processing resources), or, as proposed by the authors, that the processing component of the particular visuo-spatial task used in the study was not sufficiently complex to constrain performance.

Indeed, this suggestion is supported by the previously mentioned study by Shah and Miyake (1996), who did find span decrements in an exclusively visuo-spatial complex span task compared to a task in which the processing component did not match that used in the memory load component (verifying a sentence and remembering spatial orientation of presented items). In their study, the processing element of the visuo-spatial task involved participants having to perform mental rotation (prior to recalling spatial orientation), a task arguably more difficult than the location task utilised by Bayliss et al. (2003). Oberauer et al. (2000) also report evidence to support individual differences in the separability of spatial and verbal measures in working memory. Using a principal components analysis, Oberauer et al. found two distinct factors, each appearing to support a distinction between verbal-numerical working memory and visuo-spatial working memory. As with Shah and Miyake (1996) the visuo-spatial tasks were relatively complex (including visual tracking, spatial integration, and spatial updating), lending weight to the suggestion the tasks must transcend a particular level of difficulty in order to constrain performance.

It would seem, therefore, that a high degree of relatedness between material to be processed and stored in complex span tasks impairs complex memory span, in both children and adults (though see Towse et al., 2002). So why is complex memory span performance lower when items stored belong to the same stimulus domain as the items to be processed? According to the multiple component model of working memory (e.g., Baddeley, 1986; Duff & Logie, 2001), the cognitive demands of complex span tasks are supported by the different components working memory, such that storing the recall items in a verbal storage task would be allocated to the phonological loop, whereas performing a spatial rotation task would be carried out by the visuo-spatial sketchpad, with a possible role for the central executive in coordinating processing and storage operations. If the tasks rely on a mutual resource pool, such as in the typical reading span task, in which both the storage requirement (word recall) and the processing task (reading comprehension) rely to some extent on the phonological loop, span decrements are the result. In other words, in conditions under which two such tasks must be performed simultaneously, it is more beneficial to task performance when those tasks draw on separate systems of working memory.

However, suggestive evidence that runs contrary to this account was provided by Turner and Engle (1989, Table 1). They found that memory span performance was greatest under conditions in which the recall items (e.g., words) were unrelated to the processing material (e.g., arithmetic problems), although no statistical comparisons of the conditions were reported. Importantly, this study utilised stimuli that were both drawn from the verbal domain, as opposed to the afore-mentioned studies (Shah & Miyake, 1996; Bayliss et al., 2003), in which contrasting task elements came from the verbal and visuo-spatial domains. The working memory model cannot readily account for these findings. The phonological loop is thought to be implicated in both mental arithmetic (e.g., Logie, Gilhooly, & Wynn, 1994) and reading comprehension (e.g., Swanson, 1999; although see Gathercole & Baddeley, 1993). In complex span tasks such as those used by Turner and Engle (1989), both the processing (reading or mental arithmetic) and the storage elements (words or digits) rely more or less heavily on the phonological loop to execute the task requirements accurately. In such a case, the working memory model would predict no difference in task performance between arithmetic processing/word recall and sentence verification/word recall conditions, and arithmetic processing/word recall and arithmetic processing/digit recall conditions, as in each case, processing and storage task demands depend on the same subsystem of working memory.

To reiterate, however, the Turner and Engle (1989) findings were reported as descriptive statistics and not subjected to a statistical analysis (the study focused primarily on whether individual differences in complex span performance can predict reading comprehension ability), and as such can only be interpreted as being suggestive of an effect. However, evidence that supports the trend found by Turner and Engle comes from a study into age-related deficits in working memory by Li (1999), who found that older adults were more susceptible to similarity between numerical and verbal processing and storage stimuli than younger adults.

Taken together, the studies that have investigated stimulus-similarity effects in working memory show that memory performance is enhanced when processing and storage stimuli are dissimilar. This effect appears to occur when processing and storage material is drawn from distinct domains (verbal and visuo-spatial), but also at a more subtle level of differentiation within the verbal domain (words and numbers). However, evidence concerning the latter notion is scarce, although existing data would appear to yield reservations about the possible interpretation of these data in terms of Baddeley's (1986) working memory model. It would therefore be useful to conduct a systematic investigation using complex span tasks that crossed word and digit processing and storage requirements in order to test for stimulus-similarity decrements within the verbal domain. Moreover, with the exception of the Bayliss et al. (2003) study, stimulus-similarity effects in child populations have not been investigated to the same extent as in adults, indicating the need for further research in this area.

Two further experiments were therefore conducted to provide a systematic investigation of the effects of the similarity of processing and storage stimuli on complex span performance in seven- to ten-year old children. The experiments contrasted two types of span task (sentence span and operation span) and two different categories of the recall stimuli (words and digits). In Experiment 4, children were required to perform either a sentence completion task or a mental arithmetic task. For each span task, the children were assigned to either a Word Recall or a Number Recall condition. In the Sentence Span/Word Recall condition, the task was to recall the final words from a series of the sentences, and in the Sentence Span/Number Recall condition it was to recall a digit presented after each sentence has been processed. In the Operation Span/Number Recall condition, the task was to recall the series of arithmetic totals that had been calculated, whereas in the Operation Span/Word Recall condition, the children attempted to recall individual words presented after each arithmetic operation.

On the basis of previous evidence from studies of adults (e.g., Shah & Miyake, 1996) and children (Bayliss et al., 2003), it is predicted that memory span performance will be poorer under conditions in which the processing and recall stimuli belong to the same than different information categories. That is, the sentence span task should yield higher memory spans when the recall items are numbers than when they are words, and the operation span task should yield higher memory spans when the recall items are words rather than numbers. Children aged between seven and ten years were included – in line with relevant studies that have examined the effects of stimulus similarity – in order to investigate whether any such effects were generalisable across age groups.

3.2. Experiment 4

3.2.1. Method

3.2.1.1. Participants

A total of 96 children were drawn from Year 3 (N = 48, mean age 8 years 3 months, range 7;9 to 8;9) and Year 5 (N = 48, mean age 10 years 3 months, range 9;9 to 10;8) from a state primary school in Stockton on Tees, UK. In each age group, 28 children were randomly allocated to either the Sentence Span/Word Recall or Sentence Span/Number Recall group, and 20 children were randomly allocated to the Operation Span/Word Recall or Operation Span/Number Recall group.

3.2.1.2. Design and materials

A three-way between-subjects design was employed with type of processing task, recall category, and age as independent variables, and span as the dependent variable. The materials for the sentence span task consisted of nine sets of sentences with the final word missing (for example, *A dog wags its* _______), each set comprising two, three, or four short sentences. The interpolated memory item for the Sentence Span/Number Recall group was a single digit number (see Appendix I). The materials for the operation span task consisted of nine sets of arithmetic operations (for example, *14* + *5* = *?*), each set comprising two, three, or four equations. The interpolated memory items for the Operation Span/Word Recall group were nouns, matched in syllable with the corresponding total in the Operation Span/Number Recall group (e.g., *motorbike – twenty two*; *garden – sixteen*).

3.2.1.3. Procedure

Each participant was tested individually on a laptop computer in a quiet area of the school. All the tasks had a similar structure. In the sentence span task, participants were presented with an incomplete sentence on the computer screen, and requested to read the sentence aloud and complete it with a semantically appropriate word. No child had any problem supplying the missing word. As soon as a response had been given, the experimenter pressed a key and the sentence was replaced on the screen by a number, which was also read aloud by each child. Following another key press by the experimenter, another incomplete sentence appeared, and then another number. The children in the Word Recall group were then requested to recall, in order, the words they had generated; the children assigned to the Number Recall group were requested to recall the numbers they had read aloud. Children were presented with an increasingly long series of tasks until they failed to recall the memory items of all three trials at a particular level. Testing was discontinued at this point.

In the operation span task, the sequence of events was similar. The span task began with an arithmetic operation (e.g., *12 + 4 = ?*) displayed in black letters on a white computer screen. The children were instructed to calculate the answer as quickly and accurately as possible, and report the answer out loud. All children reached over 90 % accuracy in this task. As soon as the answer had been reported, the experimenter pressed a key and a word appeared on the screen that was to be read aloud. Following this, a further arithmetic operation appeared, followed by another word. The children in the Word Recall group were then requested to recall, in order, the words they had read aloud; the children assigned to the Number Recall group were requested to recall the totals they had calculated. As with the sentence span task, the children were presented with increasing long series of tasks until they failed to recall the memory items of all three trials at a particular level, at which point testing was terminated. A practice session preceded the task for each child. For both the operation span task and the sentence span task, each correctly recalled memory item counted as one third; the total number of thirds was then added up to provide a span score. For example, the correct recall on all the trials of one and two items, of two series of three items and two series of four items yielded a span score of $(3 + 3 + 2 + 2) \times 1/3 = 3.33$.

3.2.2. Results

Table 3 shows mean span performance for Experiment 4.

TABLE 3
Mean span performance (and standard deviations) in Experiment 4 as a function of age group, type of span task, and recall category

			Type of span task			
			Sentence span		*Operation span*	
Age group[a]	*N*	*Recall category*	*M*	*SD*	*M*	*SD*
8	48	Word	1.83	0.36	2.57	0.59
		Number	3.07	0.56	2.00	0.43
10	48	Word	2.07	0.41	3.21	0.79
		Number	3.16	0.59	2.31	0.53

[a] In years

Memory span was higher for numbers (2.55, *sd* = 0.71) than for words (2.50, *sd* = 0.78). The older children had a higher mean span (2.69, *sd* = 0.78) than the younger children (2.35, *sd* = 0.67). Spans were equivalent for the sentence completion (2.53, *sd* = 0.76) and operation span task (2.52, *sd* = 0.72).

A three-way between-subjects analysis of variance was conducted on the span scores as a function of recall category (word, number), type of span task (sentence, operation), and age (8 years, 10 years). There was no significant main effect of recall category, $F(1,88) = 3.42$, $MSe = 1.07$, $p > .05$, partial $\eta^2 = .04$, and no significant main effect of span task, $F(1,88) = 0.01$, $MSe = 0.002$, $p > .05$, partial $\eta^2 < .001$. There was a significant main effect of age, $F(1,88) = 7.76$, $MSe = 2.42$, $p < .05$, partial $\eta^2 = .08$. The span task x age interaction was nonsignificant, $F(1,88) = 1.81$, $MSe = 0.56$, $p > .05$, partial $\eta^2 = .02$, as were the recall category x age interaction $F(1,88) = 1.02$, $MSe = 0.32$,

$p > .05$, partial $\eta^2 = .01$, and the span task x age x recall category interaction, $F < 1$. However, the span task x recall category interaction was highly significant, $F(1,88) = 67.84$, $MSe = 21.14$, $p < .001$, partial $\eta^2 = .44$. The simple effects of recall category were explored for each of the span tasks using one-way between-subjects analyses of variance. In the sentence span task, number recall was significantly higher than word recall, $F(1,38) = 57.15$, $MSe = 13.58$, $p < .001$, $\eta^2 = .60$. In the operation span task, word recall was significantly higher than number recall, $F(1,54) = 18.49$, $MSe = 7.63$, $p < .001$, $\eta^2 = .26$.

3.2.3. Discussion

In this experiment, children's memory spans were superior when the stimuli encountered in the processing task and recall items were drawn from different rather than common semantic categories. In the sentence completion task, recall of unrelated numbers was greater than that of the sentence-final words generated by the participant. Similarly, in the arithmetic operation task, recall of unrelated words was superior to that of the calculated totals. The findings were consistent across both 7/8- and 9/10-year old age groups. The reversal in the operation span task of the number recall superiority established in the sentence span task rules out an account in terms simply of an intrinsic memory advantage of one stimulus category (words or numbers) over another. The findings are instead consistent with previous reports of poorer memory span performance under conditions in which the processing and recall stimuli are drawn from similar rather than distinct categories (e.g., Bayliss et al., 2003; Shah & Miyake, 2003; Turner & Engle, 1989).

However, there was in the present experiment a procedural difference between the similar and dissimilar conditions that may have contributed to the results. The stimulus-similar recall items (words in the sentence span task, digits in the operation span task) were generated directly by the processing activity. In contrast, the stimulus-dissimilar recall items (digits in the sentence span task, and words in the operation span task) were unrelated to the processing and presented subsequent to the completion of each processing activity. It is possible that the greater memory spans in the stimulus-dissimilar conditions reflects the better recall of stimuli that were unrelated to rather than directly generated by the processing activity, rather than an advantage to memory items drawn from a different domain to the processing stimuli. Related to this issue are mixed findings over whether integrated span tasks are better predictors of complex cognition than tasks in which the storage item is not generated by the processing activity (e.g., Conway et al., 2002; Süß et al., 2002; but see Turner & Engle, 1989).

However, it should be noted that there is a considerable body of evidence pointing to beneficial rather than disruptive effects of self-generation on memory performance (e.g., Slamecka & Graf, 1987). As such, this confound seems unlikely to underpin the present findings. More plausibly, the independent presentation of the memory item in the stimulus-dissimilar recall conditions may enhance temporal distinctiveness, a factor that is known to facilitate immediate memory (e.g., Neath & Crowder, 1990). In both arithmetic operation/ word recall and sentence completion/ digit recall (i.e., the stimulus-dissimilar conditions) the to-be-remembered items were presented at the end of the episode, rather than emerging as a product of the processing activity. It is possible that a recall benefit in these conditions is attributable to the temporal isolation of storage items.

A further experiment was therefore conducted with the aim of eliminating the procedural confound between the similar and dissimilar conditions in Experiment 4. In Experiment 5, the items presented for recall in both the stimulus-similar and -dissimilar conditions were unrelated to processing activities. The two tasks involved either sentence-based processing or arithmetic processing, with memory items consisting either of unrelated words or digits. In the Sentence Span/Word Recall condition, the task was to process a series of sentences for meaning, and then to recall the sequence of unrelated individual words presented after each processing activity. In the Sentence Span/Number Recall condition, single digits rather than words were presented after each sentence, for later recall. In the Operation Span/Word Recall condition, the processing task involved a series of simple arithmetic calculations, each of which was followed by the presentation of an individual word to be recalled later. In the Operation Span/Number Recall condition, single digits rather than words were presented after each calculation, for later recall. If the recall advantage of the dissimilar over similar conditions in Experiment 4 arose from differences in the manner by which the recall items were provided (generated either by the processing activity or independently by the experimenter), then differences in memory span performance across stimulus-similar and -dissimilar conditions should be eliminated in Experiment 5 as the storage items presented independently of the processing activity in both conditions. Alternatively, if recall is genuinely impaired as a result of processing and storage stimulus similarity, memory span should be greater for numbers than words in the sentence span task, and for words than numbers in the operation span task.

3.3. Experiment 5

3.3.1. Method

3.3.1.1. Participants

A total of 80 children from Year 3 ($N = 40$, mean age 7 years 7 months, range 7;0 to 8;0) and Year 5 ($N = 40$, mean age 9 years 6 months, range 9;1 to 10;0) from a state primary school in Stockton on Tees, UK. In each age group, the children were randomly allocated to one of four groups: the Sentence Span/Word Recall group, the Sentence Span/Number Recall group, the Operation Span/Word Recall group, or the Operation Span/Number Recall group. None of the children had participated in the previous experiments.

3.3.1.2. Design and materials

A three-way between-subjects design was employed with age, recall category and type of span task as independent variables, with span as the dependent variable. The materials for the sentence span task consisted of nine sets of sentences, each set comprising two, three, or four simple sentences. Thirteen sentences contained true information (e.g., *Apples grow on trees*), and fourteen sentences contained false information (e.g., *Bananas ride bicycles*). The materials for the operation span task consisted of nine sets of arithmetic operations, each set comprising two, three, or four equations (same as in

Experiment 4; see Appendix I). Items presented for retention and subsequent recall consisted either of monosyllabic nouns (e.g., *box*), or single digit numbers that ranged between 1 and 9. The numbers were generated at random, with the exception that they were never identical to the calculated total of the arithmetic task.

3.3.1.3. Procedure

Each participant was tested individually on a laptop computer in a quiet area of the school. In the reading span task, each child read a series of short sentences, and judged the veracity of each sentence in turn by responding "true" or "false". As soon as a response had been given, the experimenter pressed a key, and either an unrelated word (for the children in the Word Recall group) or a number (for the children in the Number Recall group) appeared on the screen, which remained visible until it had been read aloud by the child. This was the item which the child was instructed to retain for subsequent recall. Following a further key press by the experimenter, the next sentence appeared.

In the operation span task, the task began with an arithmetic operation (e.g., *12 + 4 = ?*) displayed in black letters on a white computer screen. The children were instructed to calculate the answer as quickly and accurately as possible, and report the answer out loud. All children reached over 90 % accuracy in this task. As soon as the answer had been reported, the experimenter pressed a key and either a word (for the Word Recall group) or an unrelated digit (for the Number Recall group) appeared on the screen that was to be read aloud. Following this, a further arithmetic operation appeared, followed by another word. At the end of each trial for both types of span task, the children in the Word Recall group were asked to recall the words, and the children in the Number Recall group were asked to recall the numbers in the order that they had been presented. This structure began with three trials of two successive tasks (i.e. [processing task – memory item], [processing task – memory item]) and increased to three trials of three tasks, then three trials of four tasks, and so on. If recall was unsuccessful on all three trials at a particular level, testing was terminated. A practice session preceded the task for each child. Span was scored as in Experiment 4.

3.3.2. Results

Table 4 shows mean span performance for Experiment 5. Memory span was higher for numbers (2.32, *sd* = 0.72) than for words (2.22, *sd* = 0.68). The older children had a higher mean span (2.48, *sd* = 0.70) than the younger children (2.06, *sd* = 0.64), and the sentence span task produced higher spans (2.41, *sd* = 0.63) than the operation span task (2.13, *sd* = 0.74).

TABLE 4
Mean span performance (and standard deviations) in Experiment 5 as a function of age group, type of span task, and recall category

			Type of span task			
			Sentence span		*Operation span*	
Age group[a]	*N*	*Recall category*	*M*	*SD*	*M*	*SD*
7	40	Word	2.00	0.44	2.00	0.85
		Number	2.53	0.28	1.73	0.64
9	40	Word	2.07	0.38	2.83	0.61
		Number	3.03	0.73	2.00	0.35

[a] In years

A three-way between-subjects analysis of variance was conducted on the span scores as a function of recall category (word, number), type of span task (sentence, operation), and age (7 years, 9 years). There was no significant main effect of recall category, $F(1,72) = 0.61$, $MSe = 0.20$, $p > .05$, partial $\eta^2 = .008$. There was a significant main effect of span task, $F(1,72) = 4.41$, $MSe = 1.42$, $p < .05$, partial $\eta^2 = .06$. There was also a significant main effect of age, $F(1,72) = 10.79$, $MSe = 3.48$, $p < .05$, partial $\eta^2 = .13$. The span task x age interaction was nonsignificant, as was the recall category x age interaction (Fs < 1). However, the span task x recall category interaction was highly significant, $F(1,72) = 26.09$, $MSe = 8.41$, $p < .001$, partial $\eta^2 = .27$. The span task x age x recall category interaction was marginally significant, $F(1,72) = 3.88$, $MSe = 1.25$, $p = .053$, partial $\eta^2 = .05$.

The simple effects of recall category were explored for each of the span tasks using one-way between-subjects analyses of variance. In the sentence span task, number recall was significantly higher than word recall, $F(1,38) = 21.60$, $MSe = 5.60$, $p < .001$, $\eta^2 = .36$. In the operation span task, word recall was significantly higher than number recall, $F(1,38) = 6.20$, $MSe = 3.01$, $p < .05$, $\eta^2 = .14$.

The simple effects of age were also explored for each of the span tasks. A one-way analysis of variance showed that there was no significant difference in memory performance between the older and younger children on the sentence span task, $F(1,38) = 2.10$, $MSe = 0.81$, $p > .05$, $\eta^2 = .05$, but older children significantly outperformed younger children on the operation span task,

$F(1,38) = 6.23$, $MSe = 3.03$, $p < .05$, $\eta^2 = .14$. In the sentence span task, younger children recalled significantly more numbers than words, $F(1,18) = 10.17$, $MSe = 1.41$, $p < .01$, $\eta^2 = .36$, and older children also recalled significantly more numbers than words, $F(1,18) = 13.80$, $MSe = 4.66$, $p < .01$, $\eta^2 = .53$. In the operation span task, older children recalled significantly more words than numbers, $F(1,18) = 13.84$, $MSe = 3.46$, $p < .01$, $\eta^2 = .43$, but there was no significant difference in memory span for the younger children, $F(1,18) = 0.63$, $MSe = 0.35$, $p > .05$, $\eta^2 = .03$.

3.3.3. Discussion

Using a procedure in which the recall items were independent of the processing task in all conditions, memory span was found to be greater for numbers than words in the sentence completion task, and conversely for words than numbers in the operation span task. These findings indicate that the corresponding pattern of findings obtained in Experiment 4 was not an artefact of the procedural differences in the manner of generation of the memory items (self-generation versus experimenter presentation) in the same-category and different-category conditions. The results suggest instead that children's performance in complex memory span tasks is genuinely impaired when the processing and recall stimuli are drawn from the same rather than different semantic categories.

Although the general pattern of similarity decrements across processing and storage domains emerged for both age groups in both experiments, some age-related differences were found in Experiment 5 on the operation span task. The similarity decrement in operation span was found only in the older age group. This asymmetry of findings was unexpected and was not reflected in the data from Experiment 4. It should, however, be noted that span scores were extremely low in the operation span / number recall condition in the younger age group (mean span 1.73). It is therefore possible that the absence of a similarity effect in operation span simply results from a floor effect in performance.

3.4. Chapter summary

Experiments 4 and 5 demonstrated that complex memory span performance in children is poorer when the type of stimuli encountered in the processing activity matched that of the items to be remembered than when the processing and storage items are drawn from different stimulus categories. This stimulus-similarity effect was extremely robust, generalising across both word- and number-based tasks, across paradigms in which the same-category items were either independent of or generated by the processing activity, and across age groups.

In summary, the experiments reported here extend previous evidence (Bayliss et al., 2003) that in children, complex span performance is influenced by the similarity between processing and storage stimuli. Contrary to reconstructive views of short-term memory, no beneficial effect of the processing context on item recall was observed: the similarity decrement occurred irrespective of whether the stimulus-similar items were the product of the processing activity, or unrelated. These findings indicate that complete accounts of working memory span will need to include mechanisms that mediate the similarity decrements, in addition to attentional constraints (e.g., Barrouillet et al., 2004) and time-based forgetting (e.g., Towse et al., 1998).

Chapter 4

Lexicality and lexical-semantic interference in working memory

---

4.1. Introduction

The previous chapter (Experiments 4 and 5) demonstrated that children's complex span performance is disrupted under conditions in which both verbal processing and storage items were either numerical or non-numerical stimuli. This finding – which will be discussed in more detail in Chapter 5 – is problematic for a multiple component working memory model (e.g., Baddeley, 1986), as the recall of verbal stimuli would be expected to be mediated by the phonological loop, regardless of whether they were digit names or not.

An alternative explanation of the findings is in terms of interference processes in working memory. Engle et al. (1999) proposed that the ability to activate and maintain memory representations in the face of interference or distraction underpins individual differences in working memory capacity. In their view, complex span tasks require controlled attention to prevent secondary information from interfering with the maintenance of target memory items. This account does not, however, explain why recall is disrupted to a greater degree by processing stimuli that are drawn from the same informational domain (word recall/ sentence processing; digit recall/ arithmetic processing). The distraction of attention is not necessarily linked to the similarity of the representations involved; thus, the Engle et al. (1999) interference account does not adequately explain the findings of a similarity-based interference effect.

A more promising explanation comes from Oberauer and colleagues (Oberauer & Kliegl, 2001; Oberauer et al., 2004; Lange & Oberauer, 2005). According to this feature overwriting account, the representations generated during the course of the processing task interfere with representations currently being maintained to fulfil the storage requirement to the extent that the generated representations (e.g., phonological, semantic, visual) share the same features or attributes (see also Saito & Miyake, 2004). However, there is little evidence that similarity between processing and storage material within a content domain disrupts complex memory span.

In order to specify more precisely the mechanisms underpinning interference in working memory, Oberauer et al. (2004) conducted a study in which the similarity within both the spatial and verbal domains was manipulated. In the spatial domain, similarity was varied by crossing two types of spatial task. In the similar conditions, the processing and storage tasks were taken from the same category (either matrix patterns or lines in a dot grid); in the dissimilar conditions, processing and storage activities were crossed (matrix processing/ line storage; line processing/ matrix processing). In the verbal domain, phonological and semantic similarity was manipulated. Recall items consisted of nouns referring to animals or plants. Participants were required to read aloud three words interleaved between the presentation of memory items. In the semantically similar conditions, the interpolated words were animate nouns (animals or plants), whereas in the low-similarity conditions, the words consisted of inanimate nouns. Phonological similarity was varied by using processing and storage words with either high or low phonological overlap. Overall, the degree of similarity between processing and storage materials was found to have little effect on performance, clearly providing a substantial challenge for feature-based interference accounts of working memory.

A contrasting account of the Oberauer et al. (2004) findings is that interference does not arise solely from feature degradation resulting from the activation of irrelevant representations. Instead, recall is largely constrained when participants cannot readily discriminate between target and non-target items at the later stage of retrieval. However, when processing and storage items are drawn from distinct or highly familiar categories (e.g., digits or words) or can be easily distinguished on the basis of modality (spatial, verbal etc.) they serve to generate cues, which in turn facilitate recall accuracy. In the Oberauer et al. (2004) study, the intrinsic features of the processing and storage items in the dissimilar conditions did not provide prominent cues with which to discriminate between target and non-target items. In the spatial conditions, all memory and processing stimuli were presented in a largely similar fashion (3 x 3 grids). Thus, there was potential for overlap between an array of features generated during the course of the task, even though the content of the grids was varied in terms of processing and storage similarity. A similar argument applies to the verbal conditions, in which target and processing items were differentiated in the semantically dissimilar condition by animacy of the nouns. Differentiation of items on the basis of animacy is neither highly familiar nor practiced in the same way as, for example, digits and words, and for this reason may not provide an effective cue for discrimination at retrieval. Finally, there was no reliable cue for selecting potential target responses in the phonologically dissimilar condition, as neither the memory nor the processing items shared any common physical features. Thus in all three cases, the manipulations of similarity between memory and processing items in the Oberauer et al. (2004) study were not implemented in a way that was likely to support easy discrimination of target items at retrieval.

The view that discriminability of target memory items can facilitate recall is consistent with findings from studies investigating intrusion errors in complex span (De Beni & Palladino, 2000). De Beni et al. (1998) compared good and poor comprehenders on verbal working memory span tasks, and found that poor comprehenders were more likely to produce intrusion errors; that is, words that had appeared within the processing phase of the span task were erroneously recalled as memory items. The authors proposed that complex span tasks rely on the capacity to inhibit irrelevant information, and that intrusion errors were a result of ineffectively disregarding or 'dumping' the information from the processing task once it had been carried out. Findings that proactive interference builds up across trials within complex span tasks (May et al., 1999; Lustig et al., 2001) provide further evidence for response confusion in working memory, in this case resulting from increases in the number of activated representations from which target memory items must be selected.

This final set of experiments provides a detailed investigation into the extent to which the disruptive consequences of similarity between memory and processing items in the verbal domain operate at phonological and lexical-semantic levels. In addition, the experiments were designed to explore whether performance decrements result from interference between representations during storage or from a failure to differentiate target from non-target representations at retrieval. Stimuli in the three experiments were monosyllabic items with a consonant-vowel-consonant structure. Memory items were either words or nonwords presented visually, and the key processing conditions involved monitoring a string of spoken words or nonwords presented between memory items for phonemic content. In Experiments 6 and 7, performance was measured using a span procedure; control conditions across these two experiments involved either articulatory suppression or no processing activity. Experiment 8 employed a fixed-list length procedure.

A feature overwriting account (Saito & Miyake, 2004; Oberauer et al., 2004) would predict a greater disruptive effect of word processing than nonword processing on the recall of words, as the representations generated by nonwords lack associated semantic attributes which could cause interference among processing and storage stimuli. In contrast, one would not expect to find a converse finding for nonword recall; that is, nonword recall should be largely unaffected by the lexical status of the processing items. This is due to the fact that nonwords do not generate semantic representations that are vulnerable to overwriting. Interference between processing and memory items at a phonological level, in line with the feature overwriting accounts of Oberauer et al. (2004) and Saito and Miyake (2004), should be equivalent for all storage and processing conditions because the degree of phonological overlap between stimuli is independent of lexical status. Recall should, however, be impaired following any phoneme monitoring activity compared with articulatory suppression interpolated between the presentation of memory items, on two grounds. First, the amount of phonological material generated during suppression (the single word "the" repeated throughout the experiment) is minimal, generating fewer phonological representations and hence a lower degree of phonological overlap with target items than the monitoring tasks. Also, the attentional demands of articulatory suppression are minimal (Barrouillet & Camos, 2001), whereas phoneme monitoring is likely to be more demanding and hence disruptive of recall.

If, however, span task performance is mediated by a failure to discriminate between target and non-target representations at retrieval, a different pattern of nonword recall performance should be observed. Nonword monitoring should impair nonword recall to a greater extent than word monitoring, as in the former processing condition, no lexical status cues are provided to allow effective discrimination of potential target from non-target items. There is independent evidence from the serial recall paradigm that lexical status (words or nonwords) is an effective cue for such discrimination (Gathercole et al., 2001), such that participants use the lexical status of the target items to guide the selection of items for output, even in error. If the same lexical consistency strategy for selecting responses can be employed in the present complex memory span task, participants should be able to differentiate potential processing from storage items under those conditions in which the lexical status of memory and processing items differs but not when they are the same. With regard to word recall, predictions generated by the feature overwriting and the discrimination cue accounts are similar: in both cases, one would expect to find greater impairment in the word processing/ word recall condition than in the nonword processing/ word recall condition. The crucial distinction between the accounts concerns the influence of word and nonword processing on nonword recall, and therefore relates to the mechanisms underpinning the stimulus-similarity effects.

A second – related – focus of this set of experiments was on the effects of lexicality of storage and processing items in working memory tasks. There appears to be very little evidence regarding the role of lexicality in complex span tasks, although in serial recall paradigms, the presence of a recall superiority for lists of words over nonwords is well-established (e.g., Hulme et al., 1991; Gathercole et al., 2001). One explanation of this lexicality effect was advanced by Hulme and colleagues (e.g., Roodenrys, Hulme, & Brown, 1993; Hulme, Roodenrys, Schweickert, Brown, Martin, & Stuart, 1997). According to this account, phonological codes of verbal items undergo rapid decay, and the availability of a representation of the phonological form of words is crucial to the retrieval process. Accurate retrieval can only occur if knowledge of the phonological structure of the items to be remembered is available (as with words), in order to reconstruct whole lexical items even if some information cannot be reconstructed from the partially decayed memory trace. This process is termed redintegration (Hulme et al., 1997; Gathercole, Frankish, Pickering,

& Peaker, 1999; Thorn, Gathercole, & Frankish, in press), and involves relatively automatic mechanisms that are thought to be an integral part of speech perception and production. According to this view, redintegration is effective for memory items with lexical representations (words), but not for those lacking such representations (nonwords). With regard to the present series of experiments, the question considered here relates to whether lexicality exerts a corresponding beneficial influence on complex memory performance. Such a finding would lend weight to accumulating evidence that serial recall and complex memory span paradigms tap some common cognitive processes (e.g., La Pointe & Engle, 1990; Lobley et al., in press).

The final issue addressed in these experiments concerns potential developmental changes in the mechanisms underpinning interference effects in working memory. Working memory function has been extensively researched in children as well as adults, with much of the theoretical analysis in the field being driven by both experimental and individual differences analyses of children's performance (e.g., Bayliss et al., 2003; Towse, Hitch, Hamilton, Peacock, & Hutton, 2005). It is worthy of note that in general, evidence points to a continuous development of fundamental cognitive abilities (Bjorklund & Harnishfeger, 1990; Case et al., 1982; Kail, 1992; Swanson, 1999). It is, however, at least possible that the use of knowledge-based cues such as lexical status to discriminate potential target from non-target responses develops across the childhood years, in which case children may be less sensitive to the lexical similarity of processing and storage items than adults. The first experiment in the present series investigates complex memory span performance in 9- and 10-

year old children. The remaining two experiments involved both child and adult participant groups, in order to test the extent to which key findings generalize across age.

4.2. Experiment 6

4.2.1. Method

4.2.1.1. Participants

Eighteen children were drawn from Year 5 of a local primary school in Stockton-on-Tees, England. They were all native English speakers and their ages ranged from 9 years 9 months to 10 years 8 months (mean age 10 years 3 months).

4.2.1.2. Design and materials

A set of 144 words and 144 nonwords, all of which had a one-syllable consonant-vowel-consonant structure, were used as processing and memory stimuli (see Appendix II). The words were taken at random from the MRC Psycholinguistic Database, with the constraint that the mean age-of-acquisition for each word was under five years (from the norms of Gilhooly & Logie, 1982). This was to ensure a high degree of lexical familiarity with the word stimuli. The nonwords were drawn from the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002). Of the 144 items in each set, 18 items had the onset phoneme /k/ (*e.g.*, *cap*, *keb*). The word and nonword sets were used to construct 42 lists for the processing task, each comprising three items. Each three-item list contained zero, one, or two items with the onset phoneme /k/, unpredictably within the list. The consonant composition of the remaining items within each list was as distinctive as possible, that is, within each processing sequence, the items contained different consonants. Each processing sequence had an associated recall item, i.e. an item that was presented at the end of the monitoring list, but was not part of the monitoring task. These recall stimuli were also drawn from the word and nonword pool, but did not include any of the items with the onset phoneme */k/*. There was no phonological overlap between recall items within a single trial.

A two-way within-subjects design was employed with type of processing activity (word processing, nonword processing, articulatory suppression) and memory item (word, nonword) as independent variables, and memory span as the dependent variable. The recall conditions were blocked; half of the participants completed the word-recall conditions first, the other half completed the nonword-recall conditions first. The order of processing activities was counterbalanced across groups of participants.

4.2.1.3. Procedure

Each child was tested individually in a quiet area of the school. The experimental stimuli were presented on a laptop computer. In the word processing and nonword processing conditions, the sequence of three processing items interpolated between memory items was presented auditorily (read aloud by the experimenter) at a rate of approximately one item per second. The memory items were presented auditorily and visually (items appeared in print on the screen and were read aloud by the experimenter). A sequence of processing items preceded the first memory item. The recall task was to remember the memory items displayed on the screen in the same order as presented. Children were also required to tap the table when they heard an item with the onset phoneme */k/* in the list of processing items.

In the articulatory suppression condition, children looked at the blank screen for three seconds while repeatedly saying the word '*the*' aloud. A metronome was set to pace the children to say one '*the*' every 750 ms. After three seconds, a memory item appeared on the screen and was read aloud by the experimenter. The children were instructed to suspend articulation while the item was on the screen. The memory item remained visible for 1 s; then the screen went blank. Again, children were requested to recall, in order, the items that had appeared on the screen. The experimenter recorded on a response sheet whether responses were correct or incorrect.

Testing began with three trials of two lists (i.e. two items for recall), followed by three trials of three lists, and so forth. The number of lists increased (to a maximum of five lists) until a child failed to recall correctly the memory items of all three trials at a particular level. Testing was discontinued at this point. Each child practiced the monitoring task, the articulatory suppression, and then one trial of processing plus recall, prior to testing.

Span was scored as follows: starting from a baseline score of one (in cases where none of the items from the two-list trials were correctly recalled), each correctly recalled memory item counted as one third; the total number of thirds was then added up to provide a span score. For example, the correct recall on all the trials of two items, of two sets of three items and two sets of four items yielded a span score of 1 + (3 + 2 + 2) x 1/3 = 3.33. Hence, the minimum score was 1.0, and the maximum score was 5.0.

4.2.2. Results

Table 5 presents the mean span scores for word and nonword recall across processing conditions. Recall of words was superior to nonwords in the nonword monitoring and articulatory suppression conditions, but not in the word monitoring condition. Word monitoring appeared to impair word recall, but nonword recall was uninfluenced by the lexicality of the processing material.

TABLE 5

Mean span scores and standard deviations for word and nonword recall across different processing activities in Experiment 6

Processing activity	*Word Recall*		*Nonword Recall*	
	M	*SD*	*M*	*SD*
Word monitoring	1.79	0.47	1.62	0.34
Nonword monitoring	3.21	0.75	1.69	0.52
Articulatory suppression	3.81	0.77	2.12	0.47
Mean	2.94		1.81	

A 2 (memory item) by 3 (processing activity) within-subjects analysis of variance was conducted on the span scores. All three terms were significant: memory item, $F(1,16)=93.20$, $MSe=34.85$, $p<.001$, partial $\eta^2 = 0.85$; type of processing, $F(2,32)=122.01$, $MSe=14.48$, $p<.001$, partial $\eta^2=0.88$; and the interaction between memory item and type of processing, $F(2,32)=27.72$, $MSe=6.12$, $p<.001$, partial $\eta^2=0.63$. Planned pairwise comparisons were conducted to explore differences in word recall and nonword recall across processing activities. Memory for words in the nonword processing condition was superior to that in the word processing condition, $t(17)=11.71, p<.05$, $d=2.44$. The articulatory suppression condition produced significantly higher word spans compared to the word processing condition, $t(17)=16.21, p<.05$, $d=0.60$, and to the nonword processing condition, $t(17)=3.84, p<.05, d=0.77$. With regard to the recall of nonwords, articulatory suppression resulted in significantly higher spans compared to the word processing condition, $t(17)=3.63, p<.05, d=1.33$, and the nonword processing condition, $t(17)=4.00$, $p<.05$, $d=0.95$. There was no significant difference in nonword recall between word and nonword processing conditions, $t(17)=0.50, p>.05, d=0.18$.

4.2.3. Discussion

There were three key findings from Experiment 6. First, recall accuracy was greater for words than for nonwords, confirming that the lexicality effect found in serial recall (e.g., Hulme et al., 1991) extends to a complex span task paradigm. Second, recall of words was impaired by word processing to a much greater degree than nonword processing, consistent with featural accounts of interference (Saito & Miyake, 2004; Oberauer et al., 2004). According to a feature overwriting hypothesis, the semantic representations generated by words during word monitoring overlap with those generated by the encoding of words as storage items, resulting in impaired recall. Third, nonword recall was disrupted to an equivalent extent by both word and nonword processing relative to articulatory suppression. As the phonological content of the articulatory suppression activity was minimal compared with the two processing conditions, this result is entirely consistent with the view that interference in working memory can result from overwriting of shared features within the phonological domain. Equally, this could reflect the increased attentional demands of the phoneme monitoring conditions relative to articulatory suppression (Barrouillet & Camos, 2001).

It should be noted that the selective effect of word processing on word recall also fits well with the notion that participants use their knowledge of the lexical status of stimuli to differentiate between target and non-target items. However, there was no corresponding decrement in nonword recall with nonword processing. If lexical status can be used to select likely target responses and reject non-target ones, nonword recall should be (but was not) disrupted most by monitoring nonword stimuli.

4.3. Experiment 7

Experiment 7 was designed to replicate the findings from Experiment 6 using a computer-controlled stimulus presentation format. This experiment also included an adult group of participants in addition to a further group of 9- and 10-year old children, in order to establish whether specific patterns of phonological and lexical interference observed in Experiment 6 with children generalises to adults. Such a developmental comparison might dissociate basic memory mechanisms from strategic ones, that is, lexical interference from the application of a lexical consistency strategy. Finally, a no-processing control condition was included in this experiment, in order to test whether suppressing articulation had a detrimental effect on span.

4.3.1. Method

4.3.1.1. Participants

Sixteen children were drawn from Year 5 of a local primary school in Stockton-on-Tees, UK. They were all native English speakers and their ages ranged from 9 years 10 months to 10 years 7 months (mean age 10 years 4 months). None of the children had participated in Experiment 6. The adult sample comprised sixteen postgraduate students, with an age range of 23 years 10 months to 44 years 3 months (mean age 27 years 2 months).

4.3.1.2. Design and materials

The processing and storage stimuli were identical to those used in Experiment 6. In this experiment, however, the task was extended to include a no-processing control condition with a list of storage items only. As in the previous experiment, a two-way within-subjects design was employed with type of processing activity (word processing, nonword processing, articulatory suppression, control) and memory item (word, nonword) as independent variables, and span as the dependent variable. The recall conditions were blocked; half of the participants completed the word-recall conditions first, the other half completed the nonword-recall conditions first. The order of processing activities was counterbalanced across participants.

4.3.1.3. Procedure

The procedure was similar to that of Experiment 6. In this experiment, however, task duration and presentation of stimuli were computer controlled. In the word processing and nonword processing conditions, participants were instructed to look at a blank computer screen while a list of three items was presented auditorily, from a recording, at a rate of one item per second. As in Experiment 6, participants were instructed to tap the table whenever a presented item had the onset phoneme */k/*. Following presentation of the final item in each set, the memory item appeared on the computer screen, and was also played aloud. The memory item remained on the screen for 1 s; then the screen went blank again. At the end of a set of lists, a question mark appeared on the screen, prompting participants to recall in serial order the items that had appeared. The articulatory suppression condition was almost identical to that in Experiment 6, except that here, the memory items were presented via an audio recording. In the control condition, participants were required to look at blank screen for three seconds, after which a memory item appeared on the screen and was presented auditorily from a recording.

4.3.2. Results

Table 6 presents the mean span scores for word and nonword recall across the different processing activities for the two age groups. In both groups, there was a sizeable recall advantage for words over nonwords in the nonword monitoring, articulatory suppression and control condition that was eliminated with word monitoring. In adults but not in children, nonword recall was impaired when the monitoring task involved nonwords rather than words.

TABLE 6

Mean span scores and standard deviations for word and nonword recall performance of adults and ten-year olds in Experiment 7

Processing activity	*Word Recall*		*Nonword Recall*	
	M	*SD*	*M*	*SD*
Adults				
Word monitoring	2.49	0.50	2.40	0.39
Nonword monitoring	3.36	0.40	2.03	0.45
Articulatory suppression	4.11	0.49	2.51	0.46
Control	5.53	0.39	3.07	0.54
Mean	3.87		2.50	
Ten-year olds				
Word monitoring	2.07	0.39	2.16	0.39
Nonword monitoring	3.16	0.75	2.03	0.42
Articulatory suppression	4.04	0.76	2.49	0.60
Control	5.22	0.51	2.75	0.63
Mean	3.62		2.36	

A 4 (processing activity) by 2 (memory item) by 2 (age group) analysis of variance was performed on the span scores. There were significant main effects of processing activity, $F(3,90) = 281.61$, $MSe = 42.58$, $p < .05$, partial $\eta^2 = 0.90$, and memory item, $F(1,30) = 280.92$, $MSe = 111.24$, $p < .05$, partial $\eta^2 = 0.90$, but not of age group, $F(1,30) = 2.99$, $MSe = 2.60$, $p > .05$, partial $\eta^2 = 0.09$. The processing activity by memory item interaction was significant, $F(3,90) = 120.59$, $MSe = 16.70$, $p < .05$, partial $\eta^2 = 0.80$, but none of the remaining interaction terms reached significance: processing activity by age, $F(3,90) = 2.23$, $MSe = 0.34$, $p > .05$, partial $\eta^2 = 0.07$, memory item by age, $F(1,30) = 0.44$, $MSe = 0.17$, $p < .05$, partial $\eta^2 = 0.01$, or processing activity by memory item by age, $F(3,90) = 0.29$, $MSe = 0.04$, $p > .05$, partial $\eta^2 = 0.01$.

Despite the absence of a significant 3-way interaction between age, processing activity and age, an a priori analysis was conducted on the span scores of each age group in order to explore differences in recall across processing activities. A 2 (memory item) by 4 (processing activity) within-subjects ANOVA was conducted on the adults' span scores. There was a significant main effect of memory item, $F(1,15) = 151.22$, $Mse = 60.08$, $p < .001$, partial $\eta^2 = 0.91$, with memory for words (3.88) superior to memory for nonwords (2.50). There was also a main effect of processing activity, $F(3,30) = 254.33$, $MSe = 21.75$, $p < .001$, partial $\eta^2 = 0.94$. The interaction between memory item and processing activity was also significant, $F(3,30) = 79.60$, $MSe = 7.67$, $p < .001$, partial $\eta^2 = 0.84$.

A set of planned pairwise contrasts were conducted to compare word and nonword recall across processing activities in the adult group. For the within-subjects *t*-tests a Bonferroni correction of $\alpha = .006$ was adopted. Memory for words in the nonword processing condition was superior to that in the word processing condition, $t(15) = 11.01$, $p < .006$, $d = 1.97$. The control condition produced significantly higher word spans than the word processing condition, $t(15) = 25.61$, $p < .006$, $d = 7.70$, the nonword processing condition, $t(15) = 19.82$, $p < .006$, $d = 5.50$, and the articulatory suppression condition, $t(15) = 12.52$, $p < .006$, $d = 3.60$. In nonword recall, the control condition yielded significantly higher spans than the articulatory suppression condition, $t(15) = 7.45$, $p < .006$, $d = 1.03$, the word processing condition, $t(15) = 6.61$, $p < .006$, $d = 1.24$, and the nonword processing condition, $t(15) = 10.16$, $p < .006$,

$d = 1.91$. In addition, there was a significant difference in nonword recall between word and nonword processing conditions, $t(15) = 3.93, p < .006, d = 0.58$, reflecting the lower levels of performance in the nonword processing condition.

A 2 (memory item) by 4 (processing activity) within-subjects ANOVA was also conducted on the children's span scores. There was a significant main effect of memory item, $F(1,15) = 121.14, Mse = 51.21, p < .001$, partial $\eta^2 = 0.89$, with memory for words (3.62) superior to memory for nonwords (2.36). There was also a main effect of processing activity, $F(3,30) = 115.35, MSe = 21.14, p < .001$, partial $\eta^2 = 0.89$. However, the main effects were mediated by a significant interaction between memory item and processing activity, $F(3,30) = 46.54, MSe = 9.07, p < .001$, partial $\eta^2 = 0.76$.

Planned pairwise contrasts between word recall and nonword recall across processing activities in the children's data (α adusted to .006, using a Bonferroni correction) revealed that memory for words in the nonword processing condition was superior to that in the word processing condition, $t(15) = 6.20, p < .006, d = 1.88$. The control condition produced significantly higher word spans than the word processing condition, $t(15) = 26.03, p < .006, d = 6.08$, the nonword processing condition, $t(15) = 10.14, p < .006, d = 3.97$, and the articulatory suppression condition, $t(15) = 5.76, p < .006, d = 2.28$. In nonword recall, the control condition yielded significantly higher spans than the articulatory suppression condition, $t(15) = 4.54, p < .006, d = 0.41$, the word processing condition, $t(15) = 3.86, p < .006, d = 0.93$, and the nonword processing condition, *t*(15) = 4.24, $p < .006$, *d* = 1.13. There was, however, no significant difference in nonword recall between word and nonword processing conditions, *t*(15) = 1.42, $p > .05$, *d* = 0.37

4.3.3. Discussion

The findings of Experiment 6 were replicated in Experiment 7: whereas children's recall of words was disrupted to a greater degree by word processing than by nonword processing, their nonword recall was impaired to an equivalent extent by word and nonword processing. However, a slightly diverging pattern of findings was obtained for the adult participants. In both word and nonword recall, adults were more generally disrupted when the processing stimuli shared the same lexical status as the memory items; that is, recall of nonwords was selectively impaired by nonword monitoring, and word recall was selectively impaired by word monitoring. Although the nonword decrement with nonword processing was not as great as the corresponding word-word decrement, it was nonetheless highly significant. In both age groups, processing led to lower levels of performance than articulatory suppression, and the no activity control condition yielded the highest spans overall. This suggests that there is a general disruptive effect of concurrent activity on span; the detrimental effect of articulatory suppression on recall as compared to the control conditions was presumably due to the fact that this activity prevented participants from using a rehearsal strategy.

These results suggest a developmental change in similarity-based interference in children and adults; however, taking into account the absence of a significant

3-way interaction between age group, processing activity and memory type, such an interpretation remains tentative. It was argued that the pattern of results obtained for the child group in Experiment 6 (and now also in Experiment 7) could readily be accommodated in terms of overwriting of overlapping semantic features of processing and storage items (Saito & Miyake, 2004; Oberauer et al., 2004). The selective impairment in nonword recall by nonword processing in the adult data, however, cannot be explained by such an account, as the degree of phonological overlap between the nonword memory items and both words and nonwords in the processing tasks was equivalent. Instead, the data are consistent with the suggestion that similarity effects result from confusion between target and non-target representations at retrieval, due to the absence of a cue to discriminate potential target from non-target responses. The data from Experiment 7 therefore suggest that the adult participants used their knowledge of the nonlexical status of the memory stimuli to distinguish word representations generated in the processing condition from the target nonwords.

There is, however, another potential reason for why the disruptive effect of nonword processing in the recall of nonwords was absent in the younger age group. Memory spans for nonword lists in the child groups in both experiments were very low: in Experiment 7, for example, the mean span score in the word processing condition was 2.16 for the children, compared with 2.40 for the adults. A potential decrement in this condition in the children's data may therefore have been masked by a floor effect. A further problem with low scores from a span procedure is that very few trials are tested in total so that, for example, an individual with a span of 2 will have been tested only on six trials. The absence of an impairment in children's nonword span scores when the processing activity involved processing nonwords may therefore have been caused by low measurement sensitivity.

4.4. Experiment 8

In order to investigate this possibility, a further experiment was conducted that compared recall of words and nonwords under conditions of either word processing or nonword processing in adults and children. Experiment 8 employed a fixed list length procedure designed to overcome the differential sensitivity of the span procedure to age and experimental conditions. Measurement sensitivity was also enhanced by increasing the number of trials tested at each individual list length from 3 to 6. All participants were tested on lists of 2, 3 and 4 recall items; adults were also tested on 5-item sequences in recognition of their greater memory spans.

One anticipated consequence of including fixed list lengths in Experiment 8 was that substantial numbers of errors would be generated, enabling a closer examination of output with regard to some of the specific predictions of the discrimination cue hypothesis. The first prediction is that error responses should have the same lexical status as the memory items. Second, if the absence of a clear cue to differentiate potential target from non-target responses underlies the poor performance in conditions in which the memory and storage items share the same lexical status, there should be an increase in the frequency of incorrect recall of items encountered during the processing activity in these conditions. Note that although these predictions do not necessarily run contrary to interference-based theories of working memory, they are central to the cue discrimination hypothesis.

4.4.1. Method

4.4.1.1. Participants

The adult sample in this experiment comprised sixteen undergraduate and postgraduate students. They were all native English speakers, and their ages ranged from 19 years 11 months to 44 years 3 months, with a mean age of 23 years 4 months. For the child sample, sixteen children were drawn from Year 5 of a local primary school in Stockton-on-Tees, UK. They were all native English speakers and their ages ranged from 9 years 10 months to 10 years 7 months (mean age 10 years 4 months).

4.4.1.2. Design and materials

The processing and memory stimuli were taken from the same pool of items as Experiments 6 and 7. The lists for the processing task contained five items (words or nonwords), of which either 0, 1 or 2 items began with the phoneme */k/.* As in the previous experiments, each list had an associated memory item that was presented at the end of the list, but was not part of the processing task. None of the memory items had the onset phoneme */k/.* The number of memory items to be recalled in serial order varied in length: six trials each of two, three and four items; for adults, the number of items to be remembered included further trials of five items. The recall conditions were blocked; half of the participants completed the word-recall conditions first, the other half completed the nonword-recall conditions first. The order of processing activities was counterbalanced across participants.

4.4.1.3. Procedure

Each participant was tested individually. Presentation of the experimental stimuli was controlled by a laptop computer. The sequence of five processing items interpolated between memory items was presented auditorily at a rate of one item per second; memory items were presented both auditorily and visually (in print on the computer screen) one second after the fifth item in each sequence of processing stimuli. A sequence of processing items preceded the first memory item. The recall task was to remember the memory items displayed on the screen in the same order as presented. Participants were also required to tap the table when they heard an item with the onset phoneme */k/* in the list of processing items. Testing began for children and adults with six trials of lists containing two memory items. List length increased by one item over successive blocks of six trials, with testing ceasing for children at list length four, and for adults at list length five. Responses were recorded manually by the experimenter at the time of testing.

4.4.2. Results

A strict serial recall criterion was adopted, according to which an item was only scored as correct if it was recalled in its original position in the sequence. Recall responses were further sub-classified into the following categories: An order error was recorded when a memory item was recalled in a different position in the list at output than at the original presentation. A memory intrusion was recorded when an item from another list in the same experimental condition was recalled. A processing intrusion was recorded when an item encountered in processing tasks was recalled; this error category was further subclassified as either a processing item from the same trial or from another trial in the same experimental condition. A novel intrusion occurred when an item that was not present in the same experimental condition was recalled; these errors were further subclassified as either word or nonword responses. The final error category was blank response, occurring when the participant did not recall any item at a particular list position.

Recall accuracy in each condition for both age groups is summarized in Table 7, which shows the mean proportion of items correctly recalled at each list length as a function of lexicality of processing item, lexicality of recall item, and list length. In both age groups, recall accuracy declined with increasing list length. The lexicality effect (superior recall of words over nonwords) was present only in the nonword processing condition for both age groups. In the word processing condition, this effect was eliminated in the child group (word and nonword recall at .48 in both conditions), and reversed in the adult group, where a higher proportion of nonwords (.46) was recalled than words (.52) when preceded by word processing. Nonword recall was lower with nonword than word processing in both age groups at all list lengths.

TABLE 7

Mean proportions (and standard deviations) of lists correctly recalled in Experiment 8 by adults and ten-year olds as a function of lexicality of processing item, lexicality of recall item, and list length

			Recall item			
			Word		*Nonword*	
Adults	*Processing item*	*List length*	*M*	*SD*	*M*	*SD*
	Word	2	.81	.13	.78	.15
		3	.46	.10	.55	.15
		4	.34	.08	.41	.08
		5	.22	.04	.32	.07
		Mean	.46		.52	
	Nonword	2	.99	.03	.71	.16
		3	.80	.05	.39	.08
		4	.64	.16	.22	.06
		5	.44	.08	.15	.04
		Mean	.72		.37	
		Total	.59		.44	
Ten-year olds	*Processing item*	*List length*	*M*	*SD*	*M*	*SD*
	Word	2	.83	.11	.83	.16
		3	.38	.12	.39	.12
		4	.23	.10	.22	.10
		Mean	.48		.48	
	Nonword	2	.95	.05	.78	.16
		3	.58	.12	.29	.14
		4	.34	.11	.16	.11
		Mean	.62		.41	
		Total	.55		.44	

A 4-way analysis of variance compared adults' and ten-year olds' accuracy scores as a function of lexicality of memory item, lexicality of processing item, and list length. Necessarily, this analysis only included data from the list lengths (2, 3 and 4) completed by both participant groups. All four main effects were significant: age, $F(1,30) = 67.95$, $MSe = 1.01$, $p<.05$, partial $\eta^2 = 0.69$; lexicality of processing item, $F(1,30) = 8.40$, $MSe = 0.15$, $p< .05$, partial $\eta^2 = 0.22$; lexicality of memory item, $F(1,30) = 97.95$, $MSe = 2.00$, $p < .05$, partial $\eta^2 = 0.77$; and list length, $F(2,60) = 434.18$, $MSe = 9.04$, $p < .05$, partial $\eta^2 = 0.94$. These terms reflect, respectively, the greater recall accuracy of the adults than the children, of memory items following nonword processing than word processing, of word than nonword lists, and of short than long sequences. Significant interactions were obtained between list length and age, $F(2,60) = 19.93$, $MSe = 0.42$, $p < .05$, partial $\eta^2 = 0.40$; between processing item and memory item, $F(1,30) = 202.33$, $MSe = 2.69$, $p < .05$, partial $\eta^2 = 0.87$; between processing item, memory item, and age, $F(1,30) = 11.80$, $MSe = 0.16$, $p < .05$, partial $\eta^2 = 0.28$; between processing item, memory item, and list length, $F(2,60) = 16.31$, $MSe = 0.13$, $p < .05$, partial $\eta^2 = 0.35$; and between processing item, memory item, list length, and age, $F(2,60) = 4.52$, $MSe = 0.04$, $p < .05$, partial $\eta^2 = 0.13$.

Further analyses were conducted to investigate detailed interactions between the lexicality of processing and memory items. Collapsed across age group, nonword recall was significantly lower with nonword than word processing, $t(31) = 6.90$, $p < .05$, $d = 2.35$, and word recall was significantly lower following word than nonword processing, $t(31) = 9.82, p < .05, d = 7.70$. In order to explore the significant 3-way interaction found between lexicality of processing and memory items with age, two further analyses were performed. First, a 2 (age) x 2 (processing item) analysis of variance on word recall was conducted, yielding a main effect of processing item, $F(1,30) = 137.79, MSe = 0.69, p < .05$, partial $\eta^2 = 0.82$ that reflected the higher scores in the nonword processing condition ($m = 0.71$) than in the word processing condition ($m = 0.51$). There was also a main effect of age, $F(1,30) = 87.84, MSe = 0.23, p < .05$, partial $\eta^2 = 0.75$, with adults recalling more ($m = 0.67$) than children ($m = 0.55$). The interaction between processing and age was significant, $F(1,30) = 14.28, MSe = 0.07, p < .05$, partial $\eta^2 = 0.32$; this was due to the superior recall of adults in the nonword processing condition only. A corresponding 2 (age) x 2 (processing item) analysis of variance performed on the nonword recall data yielded a significant main effect of processing item, $F(1, 30) = 46.96, MSe = 0.26, p < .05$, partial $\eta^2 = 0.61$, due to superior recall in the word processing condition ($m = 0.53$) than the nonword processing condition ($m = 0.40$). There was a main effect of age, $F(1,30) = 12.58, MSe = 0.11, p < .05$, partial $\eta^2 = 0.30$, with adults recalling more ($m = 0.51$) than children ($m = 0.43$) The interaction between processing and age was nonsignificant, $F < 1$.

Finally, a series of t-tests comparing nonword recall in the word and nonword processing conditions were conducted separately for each age group and list length, in order to establish under what conditions precisely disruptive effects of nonword processing on nonword recall were found. Significantly higher recall with word than nonword processing was found in the adult group at list lengths 3, 4 and 5, and for the children at list length 3 only ($p > .05$ for all remaining contrasts).

The frequency of error responses in each category as a function of condition and for each group, collapsed across all list lengths, is shown in Table 8. Levels of performance were similar across age groups, with average recall accuracy of 49.7% for children and 51.7% for adults. The most common category of error was a blank response, constituting 35.6% of all responses for children and 31.8% for adults. Order errors (migrations of memory items to non-target positions at recall) constituted 8.3% of responses for children and 11.2% for adults. Intrusion errors were less common, comprising 6.3% of errors for children and 5.3% of errors for adults. The lexical consistency between memory items and the errors responses was extremely high, at 100% of the errors sharing the same lexical status as the memory items for both children and adults under word recall conditions, 96.9% for nonword recall in adults, and 92.1% for nonword recall in children.

TABLE 8

Frequency of responses in each category in Experiment 8 for children and adults, collapsed across list lengths

	Children				*Adults*			
Recall:	*Word*		*Nonword*		*Word*		*Nonword*	
Processing:	*Word*	*Nonword*	*Word*	*Nonword*	*Word*	*Nonword*	*Word*	*Nonword*
Correct	414	536	415	354	618	968	699	497
Error:								
Blank response	235	246	348	403	431	236	466	574
Order	99	75	61	52	171	131	154	144
Intrusion:								
Memory other trial	49	6	20	16	55	7	11	43
Processing same trial	30	0	4	11	37	0	6	47
Processing other trial	28	0	2	8	24	0	6	21
Processing total	58	0	6	19	61	0	12	68
Novel word	8	1	3	7	8	2	0	2
Novel nonword	1	0	11	13	0	0	2	16
Novel total	9	1	14	20	8	2	2	18
Intrusions total	116	7	40	55	124	9	25	129
Total word errors	214	82	9	7	234	140	12	2
Total nonword errors	1	0	92	94	0	0	167	271

Note: Total responses in each condition: children 864, adults 1344

In order to compare the distributions of errors of each kind across conditions, a series of 2 (word recall, nonword recall) x 2 (word processing, nonword processing) χ^2 analyses were performed on the frequencies of each of the principal error categories, for each age group. Consider first the order errors. Their frequency did not vary across experimental conditions for either children, $\chi^2 < 1$, or adults, $\chi^2 = 1.477, p > .05, w = 0.30$, reinforcing previous findings from serial recall that the lexicality of memory items influences the accuracy of item rather than order memory (Gathercole et al., 2001).

Although the frequency of blank responses did not vary as a function of experimental condition for the children, $\chi^2 < 1$, it did for the adults, $\chi^2 = 63.956, p < .001, w = 2.0$. The latter term reflected the increased frequency of blank responses in the conditions in which the memory and processing items shared lexical status – for word as opposed to nonword processing in word recall, and for nonword compared with word processing in nonword recall. Comparisons of the distributions of errors across the two age groups established that the frequency of blank responses in word recall following word processing was significantly increased in the adults compared with the children, $\chi^2 = 28.501, p < .001, w = 0.94$, whereas there was no significant difference in blank responses across the two nonword recall conditions across age group, $\chi^2 < 1$.

The distribution of intrusion errors varied systematically across conditions in both age groups, with many more intrusion errors in word recall in the word than nonword processing conditions in both age groups. In the adult data, a corresponding increase in the frequency of intrusions errors was also apparent in nonword recall with nonword than word processing; this effect was somewhat weaker in the child data. The frequency of intrusion errors was investigated in a series of further analyses. In an initial analysis, possible differences in the frequency of total intrusion errors (collapsed across the memory, processing, and novel intrusions) across the two factors of recall (word, nonword categories) and processing (word, nonword) were explored. Significant differences were found, in both children, $\chi^2 = 71.775, p < .001$, $w = 2.12$, and adults, $\chi^2 = 169.497, p < .001, w = 3.25$. Further analyses were performed for the two age groups in each of the word recall and nonword recall conditions. No significant difference across the groups was found in the effect of the lexicality of processing material in word recall, $\chi^2 < 1$, although there was a highly significant group difference in the frequency of intrusions across the two nonword recall conditions, $\chi^2 = 20.237, p < .001, w = 0.80$. This reflects the large increase in intrusions in the nonword processing condition in the adults, but not the children. Further 2 (memory) by 2 (processing) analyses performed separately for the two age groups established highly significant differences in the distributions of both memory intrusions and processing intrusions across conditions for both groups ($p < .001$, in each case). For the memory intrusions, the increased frequency of intrusions in word recall with word than nonword processing did not differ significantly across the groups ($\chi^2 < 1$), although the increase in memory intrusions with nonword processing in nonword recall was significantly greater for adults than children

($\chi^2 = 11.843, p < .001, w = 0.61$). No significant differences across age groups were found in the corresponding analyses of the processing intrusions. Note that the novel intrusion data were not analyzed separately, due to the low frequency of this category of error.

In a final set of analyses, the frequency of error responses that were words as opposed to nonwords was compared across age groups, separately for the word recall and nonword recall conditions. The high frequency of word errors in word recall was equivalent for both groups, $\chi^2 < 1$. There was, however, a significant group difference in the nonword recall data, reflecting the greater frequency of nonword error responses in nonword recall for the adults than children, $\chi^2 = 12.060, p < .001, w = 0.61$.

4.4.3. Discussion

Experiment 8 replicated findings of substantial disruptions in word recall by word as opposed to nonword monitoring in children and adults from Experiments 6 and 7, extending their generality from a span paradigm to a fixed list length procedure that included supra-span sequence lengths. This pattern of results is consistent with proposals of interference between semantic features activated for the memory and processing items (Saito & Miyake, 2004; Oberauer et al., 2004). In this experiment, in contrast to both Experiments 6 and 7, a parallel disruptive influence of nonword monitoring on nonword recall was also found in both age groups. These data cannot readily be accounted for in terms of either lexically- or phonologically-based interference. They do, however, fit well with the hypothesis that the known lexical status of the memory items can be used as a cue to discriminate potential target from non-target responses at the time of retrieval. According to this account, processing items and other non-target stimuli cannot be easily rejected if their lexical status corresponds to that of memory stimuli, leading to increased frequency of error responses.

An important issue is why the children in Experiments 6 and 7 showed no sensitivity to the lexical status of processing material when recalling nonwords, but did so in Experiment 8. One possible source of this apparent disparity of findings is due to differences in task design. Measurement sensitivity was greatest in this final experiment, due to the employment of a fixed list length procedure with multiple list lengths that prevented scaling effects in the data and ensured equal numbers of trials in each condition. No significant impairments in nonword recall in Experiment 8 following nonword processing were found in either age group at list length 2 where performance approached ceiling levels, or at list length 4 for the children where performance levels were very low. Nonword processing deficits were, however, found at list lengths 3, 4 and 5 for adults, and at list length 3 for children.

Despite this, closer inspection of the qualitative patterns of error across nonword recall conditions in the two age groups indicates that there were genuine developmental differences in Experiment 8 that cannot readily be accounted for by scaling factors, as proportion of items correctly recalled was very similar for children and adults (50% and 52%, respectively). Some features of performance were common to both the child and adult groups. In particular, both groups showed a high degree of lexical consistency between the incorrect recall attempts and the memory items as predicted by the discrimination cue hypothesis, in both cases intrusions from both other memory and processing items were much higher when word recall was paired with word than nonword processing.

However, in adults a corresponding pattern of greater intrusion errors and blank responses was found in nonword recall following nonword than word processing. In contrast in children, the effect of nonword status of memory and processing items on the likelihood of intrusion errors was much less striking, with no substantial increase in intrusion errors in the nonword over the word processing condition. Also, the frequency with which nonword and word errors were generated in recalling nonwords in the two monitoring conditions was equivalent for children. These results indicate that the nonlexical status of nonword memory items was not as effectively used by the child group as a cue for differentiating potential target from non-target representations.

4.5. Chapter summary

Experiments 6 to 8 explored the effect of the lexical status of memory and processing stimuli on children's and adults' complex memory performance, with the aim of investigating more closely the possible mechanisms of interference in working memory. In a complex memory task, participants recalled words or nonwords while either monitoring words or nonwords for phonological content, or suppressing articulation. In 9- and 10-year old children and adults, word recall was markedly impaired by monitoring words compared with nonwords. A converse disturbance of nonword recall by nonword monitoring was consistently found for adults, but was less marked across experiments in the child groups.

Chapter 5

General Discussion

A series of eight experiments investigated the relationship between storage and concurrent processing activities in measures of working memory. The first three experiments (Chapter 3) examined the effect of different types of processing on recall performance. Experiments 4 and 5 (Chapter 4) employed sentence span and operation span tasks using either words or digits as the to-be-remembered items, in order to test whether children's complex span performance is sensitive to the similarity of processing and storage stimuli. The final series of experiments (Chapter 5) investigated the impact of the lexical status of memory and processing stimuli on memory performance in both children and adults. In this chapter, the findings from each of the experimental chapters shall be reviewed and discussed in turn, followed by a consideration of the overall findings with regard to their theoretical implications. Limitations of the study in terms of methodological and theoretical issues will be highlighted. The chapter concludes with an outline of future directions for research.

5.1. The nature of processing (Experiments 1-3)

5.1.1. Processing complexity

In Experiment 1, the processing complexity of arithmetic operations was varied under conditions in which processing times were equivalent, in order to examine whether 7- and 9-year old children's operation span performance would be affected by task difficulty. Operation span was measured with carry sums and simple sums as the processing component of the task. Contrary to a resource-sharing account of working memory (e.g. Case, 1985), children's span performance was equivalent across conditions, suggesting that the resources utilised during the processing phase of a complex span task do not draw off those resources devoted to storing the recall items. Instead, the findings from Experiment 1 lend support to the account advanced by Towse and colleagues (e.g., Towse & Hitch, 1995; Hitch et al., 2001; Towse et al., 2002), according to which children switch between processing and storage activities during the course of a complex span task. In addition, these data extend Towse and Hitch's (1995) finding that counting span is not determined by task complexity to the area of mental addition, indicating the existence of a task-general mechanism that constrains complex span performance across these two types of task.

5.1.2. Intrinsic memory demands

In Experiment 2, complex memory span of 7- and 9-year old children was assessed under three conditions designed to vary both processing and intrinsic storage demands: mental arithmetic (significant attentional demands plus intrinsic storage), odd/ even judgements (significant attentional demands, no storage required), and articulatory suppression (minimal attentional demands, no storage required). The rationale for this experiment was to examine whether the lower memory spans associated with mental arithmetic than for articulatory suppression that were observed by Barrouillet and Camos (2001; Experiment 3) were due to the intrinsic memory demands within the arithmetic processing task. The highest memory spans were found in the articulatory suppression task; span was at an intermediate level for arithmetic processing and was lowest for processing involving odd/ even judgements.

The span advantage when the interpolated task involved articulatory suppression compared with mental arithmetic replicates Barrouillet and Camos' (2001) findings, and is consistent with their view that attentionally-demanding processing activities divert limited attentional resources from storage and hence lead to accelerated temporal decay (Barrouillet et al., 2004). However, the lower levels of span performance observed in both age groups in the odd/ even than the mental arithmetic conditions were unanticipated. As both processing activities are attention-demanding, and mental arithmetic to an extent that is at the very least equivalent to and probably more demanding than the odd/even judgments, a cognitive cost account would have predicted either comparable levels of performance or an advantage to the odd/ even task (Barrouillet et al., 2004). Similarly, the decrement in odd/even span cannot be explained in terms of greater processing demands leading to reduced availability of storage according to a trade-off account (Case, 1985). In addition, the higher spans associated with the mental arithmetic task compared to the odd/ even task cannot be explained in terms of differences in intrinsic storage demands (Towse et al., 2002), as these are greater in the former than the latter tasks. Finally, the temporal equivalence of all three processing conditions rules out any account in terms of differences due to time-based forgetting (Towse & Hitch, 1995).

5.1.3. Task pacing

The equivalence of memory span in the mental arithmetic and odd/ even judgment tasks in Experiment 3 runs counter to the suggestion that lower spans associated with arithmetic processing than articulatory suppression reflect the intrinsic storage demands of the former task (Towse et al., 2002), as there is no storage burden in the odd/ even task. In Experiment 3, the pacing requirements of the interpolated processing activities for operation span and odd/ even span were equated, in order to test whether differences in task structure could account for the span differences between these two conditions observed in Experiment 2. Consistent with this suggestion, mental arithmetic and odd/ even spans did not differ, indicating that the superior recall performance in the mental arithmetic condition in Experiment 2 arose from variations in task pacing.

This explanation fits well with recent studies that have found differences in complex span performance depending on whether tasks are either self-paced or experimenter-paced (e.g., Gavens & Barrouillet, 2004; Barrouillet et al., 2004; Lépine et al., 2005). Barrouillet et al. (2004) argued that during a self-paced task, participants are free to employ different strategies to update or consolidate memory traces by postponing recall responses. When tasks are computer- or experimenter-paced, participants are forced to focus their attention on the processing task in hand, and are thereby prevented from implementing updating strategies. As a result, even relatively simple processing tasks (such as the odd/ even judgement task used here in Experiment 3) have a detrimental effect on maintenance and recall when attention switching is prevented.

Indeed, in an investigation into reading span, Friedman and Miyake (2004) found that although experimenter- and self-paced tasks were equally reliable and induced similar types of strategy use, the additional time taken to implement these strategies in the self-paced task weakened the relationship between reading comprehension and verbal SAT scores. Thus the researchers conclude with the recommendation: "*Do not allow participants to control the onset of each new stimulus, and do not allow them any time beyond that needed to process the stimuli*" (p. 155).

5.1.4. Attentional demands

Complex span in these experiments was impaired by processing activities that were attentionally demanding (mental arithmetic and odd/ even judgments), but was independent of the detailed nature of the processing involved within these activities. This pattern of findings fits well with the Barrouillet et al. (2004) view that a critical determinant of complex span is the proportion of time available to refresh item representations, and therefore that memory performance will be most impaired in tasks in which limited attentional resources have to be frequently diverted to support processing activity. Thus, span performance is mediated by the ratio of number of retrievals to units of time. The findings reported here provide support for the notion that attention must be shared between processing and storage activities. When attention is switched away from the to-be-remembered items during processing episodes, the amount of forgetting is dependent on the length of time during which recall times remain out of the focus of attention. When a task is not attentionally demanding, such as the requirement to suppress articulation, all available attentional resources can be allocated to the updating of memory items, leading to an increase in span. In conditions under which processing requirements differ in terms of complexity, but not in terms of cognitive load, there is equal opportunity to refresh memory items. Thus, mental arithmetic and odd/ even judgements produced comparable spans when the task demands were equated in Experiment 3, as participants' strategy use was constrained to an equal extent across conditions. In contrast, in Experiment 2, participants were presumably pacing the odd/ even judgements in such a way as to optimise the refreshing of memory traces during the processing episode; a strategy they were less able to implement in the mental arithmetic task.

5.2. Stimulus similarity decrements (Experiments 4 and 5)

Experiments 4 and 5 were conducted to investigate the impact of the similarity of processing and storage stimuli on children's working memory span. In Experiment 4, two types of span task were administered (sentence span and operation span), and participants were required to either recall the products of the processing task (sentence-final word, arithmetic total) or a word or digit unrelated to the processing task. Experiment 5 contrasted sentence span and operation span combined with storage of either words or digits, in tasks in which the item to be remembered was not a direct product of the processing task in either condition. In both experiments, memory span was significantly greater when the items to be recalled belonged to a different stimulus category than the material that was processed, so that in sentence span tasks, number recall was superior to word recall, and in operation span tasks, word recall was superior to number recall.

5.2.1. Self-generated v. unrelated recall items

These findings are consistent with previous studies showing span decrements with high degrees of similarity between processing and recall items in complex memory span paradigms (Turner & Engle, 1989; Shah & Miyake, 1996; Bayliss et al, 2003). This result is particularly noteworthy in Experiment 4, in which the stimulus-similar items were generated directly by the processing activity but resulted in reduced span scores. On *a priori* grounds, one might have expected a recall advantage for stimuli that have been generated directly by the processing activity over those that are unrelated to the processing activity. In episodic memory, self-generation of memory items confers a substantial benefit (Slamecka & Graf, 1978). In the specific context of this working memory task, an advantage might have been expected because memory for the processing activity provides a relevant context that could support reconstruction of degraded memory representations (e.g., Cowan, Towse, Hamilton, Saults, Elliott, Lacey, Moreno, & Hitch, 2003). On these grounds, the present finding that self-generated items were recalled more poorly than the unrelated stimuli is counter-intuitive. The findings indicate either that contextual and lexical reconstruction does not occur in complex memory span tasks or that if it does, the benefit for recall is more than offset by a disruptive effect of processing and recall items sharing the same stimulus category.

5.2.2. Trade-off between storage and processing

The crucial theoretical issue raised by the findings from Experiments 4 and 5 is why complex memory span performance is lower when items to be stored belong to the same stimulus category as items that are processed. It is unclear how such data could be handled by the notion of an undifferentiated working memory resource supporting both storage and processing (e.g., Daneman & Carpenter, 1980). If both activities are sustained by a single, generic pool of resources, there would be no reason to expect an impact of similarity of processing and storage stimuli; if anything, one might expect a recall advantage as a result of a closer association between material to be processed and recalled. The finding that stimulus similarity is detrimental to recall is therefore incompatible with a resource-sharing account. These data do not, however, rule out the possibility that resource-sharing plays a role in other working memory tasks, for example where the processing portion of the task does not prevent the use of mnemonic strategies such as grouping of items or elaborate rehearsal (e.g. Cowan, Wood, Wood, Keller, Nugent, & Keller, 1998).

5.2.3. Separate subsystems in working memory

An alternative account advanced by Duff and Logie (2001) is that the greater the separation of processing and storage demands, the more easily the information is handled by the separate subsystems of working memory such as the central executive and the phonological loop. This explanation can be readily applied to previous studies that have found stimulus-similarity decrements within verbal and visuo-spatial complex span tasks (e.g., Shah & Miyake, 1996; Bayliss et al., 2003). However, the findings from Experiments 4 and 5, that similarity within content domains also disrupts complex memory span performance, is more problematic for the multiple resources account in its current form. These experiments found poorer complex span performance in children under conditions in which both verbal storage and processing items were either numerical or non-numerical stimuli. Specifically, recall of digits was lower when the processing activity involved calculating arithmetic operations than processing the meaning of sentences. In contrast, sentence processing had a disruptive effect relative to arithmetic processing on the recall of words that were not digit names. The working memory model cannot readily accommodate the present findings, in which the contrasting stimulus categories (words and digits) are both verbal in nature, and are therefore both likely to depend on the phonological loop (e.g., Baddeley, 1986). The findings also cannot be explained in terms of differentiable working memory resource demands of handling numerical and non-numerical stimuli *per se*, as the performance decrement was greatest only when the content domains of the storage and processing stimuli were the same.

5.2.4. Similarity-based interference

A second possibility is that the detrimental effect of stimulus-similarity in complex span tasks arises from interference within working memory. According to a recent account by Saito and Miyake (2004), similarity-based interference is explained in terms of the differential degrees of representational overlap between the processing and storage stimuli. When processing and storage domains are similar, the representations generated by the processing and item maintenance activities are likely to overlap, causing interference and therefore poorer recall, than when domains are dissimilar. Furthermore, as the amount of information that must be processed increases, so does the number of representations, which increases the potential for interference and performance decrements. This account fits well with adult studies investigating the effects of susceptibility to proactive interference in span tasks (e.g., May et al., 1999; Lustig et al., 2001). According to May et al. (1999), proactive interference is likely to build up across trials within a span task, because as the set sizes become progressively larger, the competition among candidate responses also increases. Drawing the stimuli to be processed and remembered from different domains (words and numbers, in the present experiments) would therefore indeed be expected to decrease proactive interference within the span task.

5.2.5. Response competition

A final, and related, possibility is that the stimulus-similarity effect arises solely from the later response competition process. By using knowledge of the domain of the target recall items, activated representations of stimuli encountered in the processing task may be more readily rejected as potential response items under conditions in which processing and recall stimuli belong to different rather than common categories. One prediction of this account is that errors in tasks in which the recall and processing items belong to the same category should feature intrusions from the processing activity. Although insufficient errors were generated in the span procedure employed in Experiments 4 and 5 to test this prediction, such intrusion errors in complex span have been observed in other studies (De Beni & Palladino, 2000; Passolunghi & Siegel, 2001; Osaka, Nishizaki, Komori, & Osaka, 2002). A final set of experiments was therefore conducted to investigate further the potential role of response competition and representation-based interference in working memory.

5.3. Lexicality and interference (Experiments 6-8)

Experiments 6 to 8 investigated the impact of the lexical status of memory and processing stimuli on complex memory performance, with the aim of exploring possible mechanisms of interference in working memory. Overall, there were small but significant decrements in conditions in which the processing and storage items had the same lexical status, with a slightly diverging pattern for children and adult participants. Recall of words was substantially disrupted when participants monitored sequences of words rather than nonwords interpolated between memory items. Under conditions of no interpolated processing, articulatory suppression, and nonword processing, recall was superior for word than nonword lists. This result is consistent with the finding of a lexicality effect in immediate serial recall that is generally explained by the redintegrative use of primed lexical phonological representations of familiar words to fill in incomplete representations of the phonological structure of verbal memory items (e.g., Gathercole et al., 2001; Hulme et al., 1991). In these experiments, the lexicality effect was abolished when the processing activity involved nonwords. The disruptive influence of word processing on word recall occurred in both memory span (Experiment 6 and 7) and fixed list length (Experiment 8) procedures, and was present in groups of 9- and 10-year old children and adults.

5.3.1. Feature overwriting

The finding of a stimulus-similarity decrement for word recall following word monitoring is entirely consistent with feature-based theories of working memory such as those of Saito and Miyake (2004) and Oberauer and colleagues (e.g., Oberauer & Kliegl, 2001; Lange & Oberauer, 2005). According to such an account, the processing activity of a complex span task generates a variety of representations that overlap with those representations generated by the to-be-remembered items. The greater the similarity of overlapping features, the greater the extent of mutual interference, and consequently, the worse the recall performance. Encountering familiar words during the monitoring activity would be expected to activate their associated semantic features, leading to degradation in overlapping semantic representations of the items to be remembered. If a large amount of information must be processed, a greater number of overlapping representations is generated, and therefore, the greater the likelihood of subsequent interference-based forgetting. For example in the word recall/word processing condition of the final experiment, participants encountered five times as many words in the processing intervals as words to be recalled, constituting a very substantial degree of potential semantic interference.

5.3.2. Redintegration

An alternative explanation that fits well with the finding that there is no longer a superior level of recall for words over nonwords under conditions of word monitoring is that recall is critically constrained during the redintegrative process suggested to support the lexicality effect. One possibility is that the lexical representations activated by the processing stimuli can lead to false redintegration, in that a lexical stimulus that was not a memory item is incorrectly selected to reconstruct an incomplete phonological memory trace. Given the common CVC pattern shared by all memory and processing stimuli, such erroneous completions seem quite likely. In contrast, nonwords are unfamiliar and lack corresponding long-term lexical-semantic representations and cannot therefore be falsely reconstructed at output. On the basis both of lexicality effects in serial recall established in children as young as 4 years of age (Gathercole et al., 2004) and the present findings of consistent influences of word processing on word recall in children and adults, it is proposed that the cognitive process underpinning this effect represents a fundamental property of the working memory system that is present from an early age. It is also proposed that this process underpinned the selective interference between digit and non-digit stimuli across memory and processing tasks found in Experiments 4 and 5, which was found to be present in 6-year old children.

It is important to note that the key difference between these two accounts of the lexical interference effect concerns the stage at which the processing words disrupts memory performance: whereas feature-based interference accounts attribute this effect to a weakened activation of the array of features that represent the memory item, the redintegration accounts locates this effect in the subsequent process of retrieval. More detailed empirical investigations are required to distinguish between these alternative theoretical accounts.

5.3.3. Lexical cue-based discrimination

The converse finding that nonword recall is impaired when processing involves monitoring nonwords rather than words, consistently in adults and to some degree but less robustly in children, cannot be readily accommodated by either a feature-based interference account nor one based on disruption to redintegrative processes. This aspect of the results fits well instead with the hypothesis that selection of appropriate candidates for recall at the point of retrieval is facilitated by the availability of a salient cue that allows the effective discrimination of potential target from non-target items[1]. In the context of Experiments 6 to 8, the lexical status cue allows participants in the nonword recall and word processing condition to reject any representations with lexical status, and also to reject any nonlexical items in the condition in which word recall is accompanied by nonword processing. It is possible that the absence of corresponding similarity effects within content domains in Oberauer et al. (2004) may have been due to the relatively low salience and discrimination power provided in the similar conditions in this study.

In contrast to the interference process described above, it is proposed that this process is strategic rather than fundamental in nature and emerges across development. These conclusions fit well with the large body of metamemory research indicating that use of strategies for optimizing memory performance emerges relatively late during the childhood years (DeMarie & Ferron, 2003;

---

[1] This lexical cue discrimination hypothesis was developed in the context of the preparation of a manuscript submitted to the *Journal of Memory and Language*, and reflects the collaboration between the thesis author and Susan E. Gathercole.

Schneider & Pressley, 1997). The present findings indicate that in adults, word recall is constrained by a combination of the lexically-based process and cue-based selection when memory and processing stimuli share a common lexical status. In contrast, nonword recall is constrained by a single mechanism of cue-based discrimination when processing involves nonwords. In the children, the lexically-based process appears to be fully operational, but cue-based selection is somewhat less effective.

5.3.4. Intrusion errors

The notion of lexical cue discrimination can also account for several aspects of the error data from Experiment 8. First, the vast majority of substitution errors observed across all conditions matched the lexical status of the memory items, indicating that item lexicality was an important indicator for selecting a lexically appropriate, albeit incorrect, response. In conditions in which both processing and storage items shared the same lexical status (word monitoring/ word recall; nonword monitoring/ nonword recall), intrusions by non-target items were common. There were very few intrusion errors from outside the experimental stimuli (i.e., novel words and nonwords), suggesting that confusion as to which activated representations correspond to memory stimuli is the major source of error in this type of task.

Evidence that lexical status acts as a cue to discriminate potential target from non-target representations was less strong in the 9- and 10-year old children tested in these experiments than in adults. In the complex span tasks employed in Experiments 6 and 7, no significant differences in nonword recall were found in the nonword and word processing conditions. In Experiment 8, which used a more sensitive fixed list length procedure, a detrimental effect of nonword processing on nonword recall was found for children with three-item lists. However, qualitative analysis of children's nonword recall indicated that intrusions by other nonwords were not much more common in the nonword monitoring condition than in the word monitoring condition. This is in contrast both with the large numbers of intrusion errors present in the word recall word processing condition in the same group, and in the corresponding nonword conditions in the adults.

5.4. Explanations in terms of existing theoretical models

5.4.1. Resource-sharing in working memory

Overall, the findings reported here provide little support for a simple resource-sharing explanation in which working memory performance is mediated by a single flexible system fuelled by a limited capacity resource that can be flexibly allocated to support processing and storage activities (e.g. Case et al., 1982; Daneman & Carpenter, 1980; Just & Carpenter, 1992). The absence of a task difficulty effect in Experiment 1 runs counter to the suggestion that the resources available to support item retention are diminished as a consequence of the greater processing load of the more complex mental arithmetic task that involved carrying, compared to the relatively simple task of single-digit addition. Furthermore, although the span advantage for articulatory suppression in Experiment 2 can be explained by a trade-off between processing and storage resources, the decrement in odd/ even span compared to operation span cannot be accounted for by a resource-sharing account. Operation span produced more processing errors than the odd/ even task, suggesting that mental arithmetic involving carry operations was more difficult for the participants tested here than odd/ even judgements. If, as proposed by Case (1985), memory span is inversely related to processing difficulty, mental arithmetic should have yielded higher spans than the odd/ even condition, in contrast to the findings from Experiment 2.

While the remaining experiments did not provide an explicit test of the resource-sharing model, there was little evidence that processing and storage compete directly for working memory resources. The effects of the systematic manipulation of stimuli in terms of verbal/ numerical similarity and lexical status cannot be accounted for by a resource-sharing account, as the degree of similarity between processing and storage material should not affect a general capacity for resource sharing. Rather, the similarity effects obtained in Experiments 4 and 5 are consistent with suggestions of separate pools of resources that support processing and storage (e.g., Shah & Miyake, 1996; Duff & Logie, 2001). Thus, with regard to resource-sharing in working memory, the data reported in this thesis tie in with other findings that have found no evidence that individual and developmental differences in complex span reflect the differential capacity of working memory for combining processing operations and temporary storage (e.g., Towse & Hitch, 1995; Hitch et al., 2001).

5.4.2. Time-based forgetting

The results provide some evidence (Experiments 1 and 3) that favour a time-based forgetting account of children's working memory as advocated by Towse and colleagues (Towse & Hitch, 1995; Hitch et al., 2001; Towse et al., 2002), according to which children switch between processing and storage activities during the course of a complex span task. In Experiments 1 and 3, memory span performance was equivalent across different processing conditions conducted over matched durations, indicating that the time taken to perform the processing activity plays a role in children's complex span. According to a time-based forgetting account, children switch between processing and storage activities during the course of a span task. Hence, recall performance depends on the duration of the duration of the processing which in turn determines the retention period during which memory traces fade away (see also Halford et al., 1994).

However, the results do not support an interpretation based on task duration alone. Clearly, a model advocating a temporal decay explanation cannot explain the findings from Experiment 2, in which three different types of processing task performed across a set duration produced significantly different spans. In addition, a time-based forgetting account would not have predicted the observed differences in span that arose from the similarity of processing and storage material. Whereas the processing task duration was not held constant in Experiments 4 and 5, presentation of processing stimuli was computer-controlled in the final two experiments. The observed differences in span can therefore not be accounted for by processing duration alone; rather, the pattern of data requires an explanation that goes beyond a notion of time-based forgetting.

5.4.3. The multi-component working memory model

How do the findings relate to Baddeley's (1986; 2000) multi-component working memory model? According to this view, the cognitive demands of the processing task are supported by the central executive, while the temporary maintenance of recall items (and any other storage material, such as interim solutions in mental arithmetic) is supported by the phonological loop. Hence, processing and storage in working memory operate independently of one another. Differences in span should therefore reflect not only the capacity limits of the central executive, but also the capacity limits of the phonological loop and the time during which information is forgotten due to decay or output interference (e.g., Baddeley & Logie, 1999). The working memory model in its current form cannot readily accommodate the present findings. The differences in recall observed in Experiment 2 – and especially the span advantage for the articulatory suppression condition – cannot be attributed to the handling of storage material by the phonological loop (and nothing else). The articulatory suppression condition should have suffered to the same extent as the mental arithmetic and odd/ even conditions, due to the prevention of active sub-vocal rehearsal of items by the repetition of the word "the". The finding that articulatory suppression produced the best recall performance indicates that other mechanisms, specifically those related to the processing activity, also constrain complex span performance.

A multi-component view is also problematic in explaining the findings from the experiments investigating stimulus-similarity decrements. This model readily accommodates previously reported similarity effects in simple span (e.g., Baddeley, 1966), in that similar items may produce phonologically-based confusion effects at output, and complex span (e.g., Bayliss et al., 2003), when the contrasting processing and storage items are drawn from verbal and visuo-spatial domains. However, in Experiments 4 and 5, the contrasting stimulus categories (words and digits) were both verbal in nature, and as such are therefore both likely to depend on the phonological loop (e.g., Baddeley, 1986). The findings also cannot be explained in terms of separate working memory resources for numerical and non-numerical stimuli, as the greatest performance decrements were observed only when the content domains of the storage and processing stimuli were the same. While the present data do not rule out the notion of separable resources in working memory, it is clear that an interpretation is necessary that incorporates a role for the interrelationship between processing and storage.

5.4.4. Attention and working memory

A more promising explanation for the pattern of findings presented here lies in an account of attentional resources (e.g., Engle et al., 1999; Barrouillet et al., 2004; Gavens & Barrouillet, 2004). According to the account proposed by Engle and colleagues (Engle et al., 1991; Conway & Engle, 1996; Engle, Conway, Tuholski, & Shisler, 1995), working memory capacity reflects the ability to activate memory representations, bringing them into the focus of attention and holding them there. The amount of attentional activation available to each individual varies, and is limited to a relatively small number of memory representations. Hence, individual differences in working memory capacity are a result of variations in attentional capacity. During the course of a complex span task, the to-be-recalled memory items receive activation from controlled attentional focusing as long as they remain within the focus of attention; if they no longer receive attention, the memory items decay (Cowan, 1995). Consequently, participants must switch their attention rapidly between the processing portion of the task and the decaying memory representations.

A related account of attentional resource-sharing was advanced by Barrouillet and colleagues (Barrouillet & Camos, 2001; Barrouillet et al., 2004). According to this account, the critical factor underlying performance on complex span tasks is the extent to which the processing task captures attention over a set period of time. As such, this model shares similarities with models of working memory that conceive of resources as a kind of 'mental energy' available for activation (Anderson, 1993; Anderson & Lebiere, 1998). Both the controlled attention and cognitive load accounts can offer an explanation for a number of findings reported here. In Experiments 1 and 3, no differences in span were found across conditions that varied in terms of the nature of the processing activity. While the processing portions of the task differed in terms of difficulty (carry operations v. simple arithmetic, Experiment1; mental arithmetic v. odd/ even judgements, Experiment 3), it can be argued that the contrasting conditions required activation from attentional focusing to a comparable extent, given that the task durations were held constant across conditions.

Of course, as acknowledged earlier, these findings can also be accounted for by a time-based forgetting model (e.g., Towse & Hitch, 1995). However, the findings from Experiment 2 are crucial in distinguishing between the time-based and attentional resource models. In this experiment, recall performance was not determined by task difficulty (which varied across conditions) or task duration (held constant). Instead, an interpretation of results in terms of attention is that the span scores differed as a result of differences in attentional requirements: articulatory suppression requires no attention, whereas mental arithmetic and odd/ even judgments do. The span advantage for mental arithmetic can be explained by the fact that this condition (serving as the temporal guide for the other two conditions) was self-paced. Thus, children had the opportunity to use portions of the processing time to refresh representations of the to-be-recalled storage items. In the odd/ even condition, there was no such opportunity, as presentation of stimuli was at a set pace. The rapid switching of attention between processing and storage was therefore impeded, leading to greater forgetting. Hence, the results from Experiments 1-3 are consistent with the view that complex span performance in children is disrupted by processing activities that divert attentional resources from storage.

While similar to the Engle et al. (e.g., 1999) account in terms of attention-switching, it is important to note that Barrouillet et al. (e.g., 2004) favour the notion of time-based decay of activation over the idea of loss of activation due to interference processes. This is in contrast to Engle et al., who emphasise that the maintenance and retrieval of activated memory representations are vulnerable to distracting events under conditions of interference. This distinction is important, as the findings from Experiments 4 to 8 require an explanation of recall decrements not only in terms of temporal decay, but also in terms of competing memory representations.

5.4.5. Representation-based forgetting

While Engle et al. (1999) emphasise a role for interference in explaining working memory decrements under certain conditions, their view is based on interference through the distraction of attention. According to this view, some of the limited attentional resources are diverted from the primary task by representations that are irrelevant to it. Oberauer and colleagues (Oberauer & Süß, 2000; Oberauer & Kliegl, 2001; Lange & Oberauer, 2005), however, argue that interference occurs through partial overwriting of overlapping representations (see also Saito and Miyake, 2004). Specifically, the processing of information during the course of a complex span task will result in the activation of a variety of representations (phonological, lexical, semantic etc.). If the representations generated by the storage requirements of the task share features with those generated by the processing activity, the overlap can lead to contamination or loss of the original representations (see also Nairne, 1990).

The crucial difference between attentional-based interference and representation-based interference is the impact of similarity between items on recall. According to the attentional-based interference view, controlled attention is used to prevent distracting secondary information from interfering with the maintenance of target memory items. This account does not, however, explain why the greatest disruptions in memory performance arise when the processing and storage stimuli are drawn from common representational domains. The results from Experiment 4 and 5 are therefore more consistent with proposals of interference between features activated for the memory and processing items (Saito & Miyake, 2004; Oberauer & Kliegl, 2001), and also fit well with the ACT-R computational model (Anderson & Lebiere, 1998; Lovett, Reder, & Lebiere, 1999), according to which confusions involving the retrieval of one node for another will tend to be limited to nodes of the same structure (for example, misretrieving one word for another).

A representation-based interference account can also be applied to certain aspects of the final set of data. In Experiments 6 to 8, recall of words was substantially disrupted when participants monitored sequences of words rather than nonwords interpolated between memory items. This finding is consistent with featural accounts of interference (Saito & Miyake, 2004; Oberauer et al., 2004), according to which overlap between the semantic features of words encountered as memory items and as processing items will lead to loss of information in the word processing but not the nonword processing condition. In Experiment 6, children's nonword recall was disrupted to an equivalent extent by both word and nonword processing relative to articulatory suppression. As the phonological content of the articulatory suppression activity was minimal compared with the two processing conditions, this result is entirely consistent with the view that interference in working memory can result from overwriting of shared features within the phonological domain. Equally, this could reflect the increased attentional demands of the phoneme monitoring conditions relative to articulatory suppression (Barrouillet & Camos, 2001). However, the finding in Experiments 7 and 8 that nonword recall is impaired when processing involves monitoring nonwords rather than words is not readily accommodated by a feature-based interference account, and is more in line with the mechanism of cue-based retrieval described above.

Thus, the pattern of overall findings can – in the main – be accounted for by existing theoretical models, and suggests that the effects of task duration, attentional demands, and interference mechanisms all play a role in complex working memory span.

5.5. Developmental considerations

This section concerns potential developmental changes in the mechanisms underpinning interference effects in working memory. It is notable that by and large, research has established developmental continuity rather than discontinuities, with children's working memory performance showing similar influences of key variables to that of adults (Bjorklund & Harnishfeger, 1990; Case et al., 1982; Kail, 1992; Swanson, 1999). This reflects the general pattern of findings reported in this thesis. Overall, little evidence was found to support the notion of qualitative developmental change in working memory performance. The absence of significant age-related interactions in Experiments 1 and 2 provides no support for such change in the mechanisms underpinning complex span performance at these ages. At both ages 7 and 9, processing activities that imposed significant processing demands resulted in lower span scores than an undemanding processing task, articulatory suppression, despite temporal equivalence of the processing conditions. Similarly, the findings from Experiments 4 and 5 suggest that the stimulus-similarity effect observed here generalises across age groups, in line with previous studies using participants of comparable ages (e.g., Bayliss et al., 2003).

The most interesting results relating to developmental change were found in the final set of experiments that investigated lexical-semantic interference in working memory. In both children and adults, word recall was markedly impaired by monitoring words compared with nonwords. A converse disturbance of nonword recall by nonword monitoring was consistently found for adults, but was either absent or less marked across experiments in the child groups. However, low measurement sensitivity may have been the cause of the absence of an impairment in children's nonword span scores in Experiments 6 and 7, and may also have contributed to the reduced effects of nonword processing on nonword recall in the adult group in Experiment 7. The findings suggest that whereas the lexical-semantic processes of either interference between memory and processing stimuli or redintegration appears to be invariant with age, the strategic use of lexical status to discriminate potential target from non-target items appears to be robust in adults but in the early stages of emergence with the younger participants.

The findings therefore provide some evidence that the use of knowledge-based cues such as lexical status to discriminate potential target from non-target responses develops across the childhood years.

5.6. Future directions

5.6.1. Maintenance or retrieval?

With regard to the interference effects observed in Experiments 6 to 8, the question remains as to whether the rehearsal set itself is influenced by interference, or whether confusion occurs during the subsequent process of retrieval. According to a confusion account, one item is confused with another, and similarity increases the competition of activated items at recall. Interference during the course of item maintenance is though to occur through feature overwriting, which leads to degradation of some features. More detailed empirical investigation is needed to disentangle these alternative accounts, although some evidence exists at least in serial recall to support the feature overwriting account (Lange & Oberauer, 2005).

5.6.2. Lexicality and working memory

The findings from Experiments 6 to 8 indicate that the lexicality effect is equivalent in both complex memory and serial recall paradigms, in both cases exerting a beneficial influence on memory performance, suggesting that a common redintegration process is applied in both cases. The current findings lend weight to accumulating evidence that serial recall and complex memory span paradigms tap some common cognitive processes (LaPointe & Engle, 1990; Lobley et al., in press); thus, the role of lexicality in working memory may prove a fertile area for future investigation.

5.6.3. Strategy use in complex span tasks

The absence of empirical support for a unitary, resource-sharing view of working memory (e.g., Case, 1985) does not rule out the possibility that resource-sharing occurs under other circumstances. For example, one potential area of investigation is whether resource-sharing plays a role in other working memory tasks, for example, when the processing portion of the task does not prevent the use of mnemonic strategies such as grouping of items and elaborate rehearsal (e.g., Cowan et al., 1998). Furthermore, the present data do not address directly the use of strategies in task-switching, for example, whether it can prevent the implementation of idiosyncratic strategies such as rehearsal.

5.6.4. Intrinsic memory demands

While there was little evidence to suggest that intrinsic memory demands affect span performance (Experiment 2), more direct experimental manipulations of memory demands of processing activities are needed to provide stronger tests of the hypothesis that item storage in complex span paradigms is influenced by the storage demands of processing activities, particularly as previous work suggests that intrinsic memory demands in some types of mental arithmetic can affect children's working memory (Adams & Hitch, 1997).

5.6.5. A role for semantic short-term memory in complex span tasks?

The findings from this thesis have shown that complex memory span appears to be a multifaceted phenomenon drawing on many levels of representations. The findings from Experiments 6 to 8 suggest that one type of representation activated during working memory tasks relates to temporarily maintained semantic information. This, then, begs the question of where to locate a short-term store for semantic representations within a working memory model. The lexical-semantic knowledge accessed by familiar words cannot plausibly be located the temporary storage capacities of Baddeley's (1986) multi-component working memory, as the slave systems are hypothesised to handle only the short-term storage of verbal and visuo-spatial information. However, their activated features may well be conceived as part of a working memory system. Disruptive effects of semantic similarity of target and distractor stimuli have been established in a variety of paradigms, including picture naming (Damian & Bowers, 2003; Vigliocco, Vinson, & Siri, 2005), word naming (Colangelo, Buchanan, & Westbury, 2004), and cross-language translation (Bloem, van den Boogaard, & La Heij, 2004).

As mentioned in Chapter 1, Martin and colleagues (Martin & Romani, 1994; Martin et al., 1994; Romani & Martin, 1999) have proposed – on the basis of cognitive and neuropsychological dissociations – that semantic traces are maintained temporarily within a component of short-term memory associated with the prefrontal cortex. While the phonological loop stores phonologically decaying traces that are refreshed through subvocal rehearsal, semantic short-term memory is hypothesised to store lexical-semantic representations (i.e., word meanings) that are actively maintained until they can be integrated into the task in hand (Haarman et al., 2001). For example, Cowan et al. (2003) recently suggested that children may use the semantic or lexical context of a reading span task as a context for retrieval. This was demonstrated by longer response times in reading and listening span as compared to operation span. In a study of adults' complex span performance, Haarman et al. (2001) reported that comprehension and verbal problem solving were better predicted by 'conceptual span', a semantic-based complex span task, than simple span measures. They suggest that the reason why complex span is a better predictor of cognitive ability than simple span may be due to the failure of simple span measures to sufficiently engage semantic STM. Thus, while the present studies provide no direct clues regarding the existence of a short-term memory dedicated to the storage of semantic information, the question of whether activated short-term conceptual representations contribute to performance on working memory tasks as well as other language–related activities merits further investigation.

5.7. Summary and conclusions

The findings presented here illustrate that complex span performance in children and adults is mediated by a constellation of factors: both the way in which complex span tasks are combined, the features of processing and storage items, as well as specific task requirements such as processing duration, are important in shaping performance. On the basis of the experimental findings in this thesis it is proposed here that working memory – and consequently performance on working memory span tasks – can be limited by four main factors. The first factor relates to the duration of the processing component of span tasks. Memory traces decay over time unless they are actively maintained by focused attention or strategies such as rehearsal. This temporal dimension of complex memory span is crucial for (at least) two reasons with regard to development: Firstly, processing speed increases throughout the childhood years (e.g., Kail, 1992; Kail & Salthouse, 1994; Fry & Hale, 1996), enabling older children to complete the processing portion of the task faster than younger children, thereby allowing them to spend more time refreshing decaying memory items. Secondly, younger children are less skilled in the use of strategies such as chunking and rehearsal than older children and adults (e.g., Flavell, Beach, & Chinsky, 1966; Gathercole, Adams, & Hitch, 1994), resulting in the greater probability of decay of the to-be-recalled items.

The second factor likely to constrain working memory performance is limits in the amount of available attentional resources. During the course of a complex span task, attentional resources are required not just to complete the processing task, but also to refresh the to-be-remembered items. Individuals with greater attentional energy are less affected by the cognitive load of the processing task (in other words, they are less constrained by time parameters) and are also able to maintain more items in the focus of attention at any given time. It is likely that such an attentional capacity increases with age, leading to higher working memory spans for older children and adults. In addition, some have argued that adults can apply higher levels of activation to items than children (Gavens & Barrouillet, 2004).

Related to the second factor is the notion of task- or attention-switching efficiency. In order to perform a complex span task successfully, participants must frequently switch between processing and storage in order to reactivate decaying memory traces. The more attentionally demanding the processing activity, the greater the impact on the maintenance of recall items. This ability has been proposed as an executive control function (Miyake et al., 2000), and involves the disengagement of an irrelevant task set and the subsequent active engagement of a relevant task set. This switching of attention is constrained by the abilities of a limited-capacity control system such as Baddeley's (1996) central executive, which is thought to be subject to developmental increase in functional capacity throughout childhood (e.g., Gathercole et al., 2004).

The final factor underpinning working memory performance relates to the ability to resist interference. This ability is characterised by two distinct cognitive processes: the ability to suppress activated irrelevant representations, a process proposed to represent a fundamental property of the working memory system that is present from an early age, and the strategic process of cue discrimination, which emerges across development. Developmental differences in span performance can therefore – at least in part – be explained in terms of strategy acquisition, and resistance to interference, a factor known to possess a steep developmental trajectory (e.g., Bjorklund & Harnishfeger, 1990).

It is important to note that – in line with the explanations advanced here – there is now a general consensus that no single factor constrains complex memory span. Experimental research over the past decade designed to isolate specific processes in working memory has generated many apparently contradictory findings that challenge the credibility of simple conceptualizations of working memory constraints (Barrouillet et al., 2004; Saito & Miyake, 2004; Towse et al., 2005). As such, the explanations provided here do not assume to provide a complete account of working memory performance. Instead, this thesis has confirmed the notion that working memory span is a complex phenomenon drawing on many levels of representation. Working memory performance in children and adults appears to be the result of a variety of factors, including the potential for interference, use of strategies, and memory demands.

Consequently, while providing new insights into the nature of the relationship between storage and processing in complex span tasks, the thesis has also highlighted important theoretical and empirical issues that merit further systematic investigation.

References

Adams, J.W. & Hitch, G.J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology, 67*, 21-38.

Anderson, J.R. (1993). *Rules of the Mind.* Hilldale, NJ: Erlbaum.

Anderson, M.C., & Neely, J.H. (1996). Interference and inhibition in memory retrieval. In E.L. Bjork & R.A. Bjork (Eds.), *Memory. Handbook of perception and cognition* (2nd ed., pp. 237-313). San Diego, CA: Academic Press.

Anderson, J.R., Reder, L.M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology, 30*, 221-256.

Anderson, J.R., & Lebiere, C. (1998). *Atomic components of thought.* Hillsdale, NJ: Erlbaum.

Andrade, J. (2001). *Working Memory in Perspective.* Hove, UK: Psychology Press.

Awh, E., Jonides, J., Smith, E.E., Schumacher, E.H., Koeppe, R.A., & Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. *Psychological Science, 7*, 25-31.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory Vol. 2* (pp. 89-195). New York: Academic Press.

Baddeley, A.D. (1966). 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. (1986). *Working Memory.* Oxford: OUP.

Baddeley, A.D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A*, 5-28.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417-423.

Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G.A. Bower (Ed.), *Recent Advances in Learning and Motivation, Vol. 8.* New York: Academic Press.

Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589.

Baddeley, A.D., Lewis, V.J., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology, 36A*, 233-252.

Baddeley, A.D., Vallar, G., & Wilson, B.A. (1987). Sentence comprehension and phonological memory: Some neuropsychological evidence. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp. 509-529). Hove, UK: Lawrence Erlbaum Associates.

Baddeley, A.D., Bressi, S., Della Sala, S., Logie, R.H., & Spinnler, H. (1991). The decline of working memory in Alzheimer's-Disease: A longitudinal study. *Brain, 114*, 2521-2542.

Baddeley, A.D., & Logie, R.H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.) *Models of Working*

Memory: Mechanisms of Active Maintenance and Executive Control (pp. 28-61). Cambridge: University Press.

Baddeley, A.D., & Wilson, B.A. (2002). Prose recall and amnesia: Implications for the structure of working memory. *Neuropsychologica, 40,* 1737-1743.

Bayliss, D.M., Jarrold, C., Gunn, MD., & Baddeley, A.D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General, 132,* 71-92.

Bayliss, D.M., Jarrold, C., Baddeley, A.D. & Gunn, MD. (2005). The relationship between short-term memory and working memory: Complex span made simple? *Memory, 13,* 414-421.

Barrouillet, P. & Camos, V. (2001). Developmental increase in working memory span: resource sharing or temporal decay? *Journal of Memory and Language, 45,* 1-20.

Barrouillet, P., Bernadin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133,* 83-100.

Bjorklund, D.G., & Harnishfeger, K.K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inhibition. *Developmental Review, 10,* 48-71.

Bloem, I., van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language, 51,* 307-323.

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology, 26*, 467-482.

Bruyer, R., & Scailquin, J.-C. (1998). The visuo-spatial sketchpad for mental images: Testing the multicomponent model of working memory. *Acta Psychologica, 98*, 17-36.

Camos, V., Fayol, M., & Barrouillet (1999). Counting in children: Double task or procedure. *L'Année Psychologique, 99*, 623-645.

Carter, C.S., Mintun, M., & Cohen, J.D. (1995). Interference and facilitation effects during selective attention: An $H_2{}^{15}O$ PET study of stroop task performance. *Neuroimage, 2*, 264-272.

Case, R. (1985). *Intellectual Development: Birth to Adulthood.* New York: Academic Press.

Case, R., Kurland, M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386-404.

Chi, M. (1978). Knowledge structures and memory development. In R. Siegler (Ed.), *Children's Thinking: What develops?* Hillsdale, NJ: Erlbaum.

Cocchini, G., Logie, R.H., Della Sala, S., MacPherson, S.E., & Baddeley, A.D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition, 30*, 1086-1095.

Colangelo, A., Buchanan, L., & Westbury, C. (2004). Deep dyslexia and semantic errors: A test of the failure of inhibition hypothesis using a semantic blocking paradigm. *Brain and Cognition, 54*, 232-234.

Conrad, R., & Hull, A.J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429-432.

Conway, A.R.A., & Engle, R.W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*, 450-466.

Conway, A.R.A., & Engle, R.W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory, 4*, 577-590.

Conway, A.R.A, Cowan, N., Bunting, M.F., Therriault, D.J., & Minkoff, S.R.B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163-183.

Conway, A.R.A., Kane, M.J., Bunting, M.F., Hambrick, D.Z., Wilhelm, O., & Engle, R.W. (in press). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review.*

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104*, 163-191.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96*, 341-370.

Cowan, N. (1993). Activation, attention, and short-term memory. *Memory and Cognition, 21*, 162-167.

Cowan, N. (1995). *Attention and memory: An integrated framework.* Oxford: University Press.

Cowan, N, Wood, N.L., Wood, P.K., Keller, T.A., Nugent, L., D., & Keller, C.V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General, 127*, 141-160.

Cowan, N., Towse, J.N., Hamilton, Z., Saults, J.S., Elliott, E.M., Lacey, J.F., Moreno, M.V., & Hitch, G.J. (2003). Children's working memory processes: A response-timing analysis. *Journal of Experimental Psychology: General, 132*(1), 113-132.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Experimental Psychology, 86*, 77-82.

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review, 10*, 111-117.

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

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

DeBeni, R., Palladino, P., Pazzaglia, F., & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *Quarterly Journal of Experimental Psychology, 51*, 305-320.

DeBeni, R. & Palladino, P. (2000). Intrusion errors in working memory tasks: Are they related to reading comprehension ability? *Learning and Individual Differences, 12*, 131-143.

DeMarie, D., & Ferron, J. (2003). Capacity, strategies, and metamemory: Tests of a three-factor model of memory development. *Journal of Experimental Child Psychology, 84*, 167-193.

Duff, S.C., & Logie, R.H. (1999). Storage and processing in visuo-spatial working memory. *Scandinavian Journal of Psychology, 40*, 251-259.

Duff, S.C. & Logie, R.H. (2001). Processing and storage in working memory span. *Quarterly Journal of Experimental Psychology, 54*A (1), 31-48.

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

Engle, R.W., Conway, A.R.A., Tuholski, S.W., & Shisler, R.J. (1995). A resource account of inhibition. *Psychological Science, 6*, 122-125.

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

Flavell, J., Beach, D., & Chinsky, J. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37*, 283-299.

Friedman, N.P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language, 51*, 136-158.

Fry, A.F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science, 7*, 237-241.

Fuster, J.M. (1997). *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe.* New York: Raven.

Gathercole, S.E., & Baddeley, A.D. (1989). 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. (1993). *Working Memory and Language.* Hove, UK: Erlbaum.

Gathercole, S.E., & Hitch, G.J. (1993). Developmental changes in short-term memory: A revised working memory perspective. In A. Collins, S.E. Gathercole, M.A. Conway, & P.E. Morris (Eds.), *Theories of memory* (pp. 189-210). Hove, UK: Erlbaum.

Gathercole, S.E., Adams, A.M., & Hitch, G.J. (1994). Do young children rehearse? An individual differences analysis. *Memory & Cognition, 22*, 201-207.

Gathercole, S.E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 25*, 84-95.

Gathercole, S.E., Pickering, S.J., Hall, M., & Peaker, S.M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology A, 54*, 1-30.

Gathercole, S.E., Pickering, S.J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental*

Psychology, 40, 117-190.

Gathercole, S.E., Pickering, S.J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*, 1-16.

Geary, D.C., Hoard, M.K., Byrd-Craven, J., & DeSoto, M.C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology, 88*, 121-151.

Gavens, N., & Barrouillet, P. (2004). Delays of retention, processing efficiency, and attentional resources in working memory span development. *Journal of Memory and Language, 51*, 644-657.

Gelman, R., & Meck, E. (1983). Preschooler's counting: Principles before skills. *Cognition, 13*, 343-359.

Gilhooly, K.J. & Logie, R.H. (1982). Word age-of-acquisition and lexical decision-making. *Acta Psychologica, 50*, 21-34.

Guyer, B.L., & Friedman, M.P. (1975). Hemispheric processing and cognitive styles in learning-disabled and normal children. *Child Development, 46*, 658-668.

Haarmann, H.J., Davelaar, E.J., & Usher, M. (2003). Individual differences in semantic short-term memory capacity and reading comprehension. *Journal of Memory and Language, 48*, 320-345.

Halford, G.S., Maybery, M.T., O'Hare, A.W., & Grant, P. (1994). The development of working memory capacity. *Child Development, 65*, 1338-1356.

Hanten, G., & Martin, R.C. (2000). Contributions of phonological and semantic short-term memory to sentence processing: Evidence from two cases of closed head injury in children. *Journal of Memory and Language, 43*, 335-361.

Hitch, G.J., Towse, J.N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General, 130*(2), 184-198.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology, 38*, 241-253.

Hulme, C., Maughan, S., & Brown, G.D.A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*, 685-701.

Hulme, C.M., Roodenrys, S., Schweickert, R., Brown, G.D.A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23*, 1217-1232.

Hutton, U.M.Z., & Towse, J.N. (2001). Short-term memory and working memory as indices of children's cognitive skills. *Memory, 9*, 383-394.

Just, M.A. & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 536.

Kail, R. (1992). Processing speed, speech rate, and memory. *Developmental Psychology, 28*, 899-904.

Kail, R., & Salthouse, T.A. (1994). Processing speed as a mental capacity. *Acta Psychologica, 86*, 199-225.

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

Lange, E., & Oberauer, K. (2005). Overwriting of phonemic features in serial recall. *Memory, 13*, 333-339.

LaPointe, 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.

Leather, C.V., & Henry, L.A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. *Journal of Experimental Child Psychology, 58*, 88-111.

Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high-level cognition. *Psychonomic Bulletin & Review, 12*, 165-170

Li, K.Z.H. (1999). Selection from working memory: On the relationship between processing and storage components. *Aging, Neuropsychology, and Cognition, 6*, 99-116.

Lobley, K., Gathercole, S. E., & Baddeley, A. D. (in press). Is there a role for the phonological loop in verbal complex span? *Quarterly Journal of Experimental Psychology.*

Logie, R.H. (1995). *Visuo-Spatial Working Memory.* Hove, UK: Lawrence Erlbaum Associates.

Logie, R.H., Zucco, G., & Baddeley, A.D. (1990). Interference with visual short-term memory. *Acta Psychologica, 75*, 55-74.

Logie, R.H., Gilhooly, K., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition, 22*, 395-410.

Logie, R.H., Della Sala, S., Wynn, V., & Baddeley, A.D. (2000). Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology, 53A*, 626-646.

Lovett, M.C., & Anderson, J.R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology, 31*, 168-217.

Lustig, C., May, C.P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology General, 130*, 199-207.

McGeoch, J.A., & MacDonald, W.T. (1931). Meaningful relation and retroactive inhibition. *American Journal of Psychology, 43*, 579-588.

Martin, R.C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology, 9*, 506-523.

Martin, R.C., Shelton, J.R., & Yaffee, L.S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language, 33*, 83-111.

Martin, R.C., & Freedman, M. (2001). Verbal working memory: The ins and outs of phonological and lexical-semantic retention. In H.L. Roediger, J.S. Nairne, I. Neath, & A.M. Suprenant (Eds.), *The Nature of Remembering: Essays in Honor of Robert G. Crowder.* Washington, D.C.: American Psychological Association

May, C.P., Hasher, L., & Kane, M.J. (1999). The role of interference in memory span. *Memory & Cognition, 27*, 759-767.

Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A.Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 442-481). New York: Cambridge University Press.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., & Wagner, T.D. (2000). The unity and diversity of executive functions and their contribution to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100.

Nairne, J.S. (1990). A feature model of immediate memory. *Memory & Cognition, 18*, 251-269.

Neath, I., & Crowder, R.G. (1990). Schedules of presentation and temporal distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 316-327.

Oberauer, K., Süß, H.-M., Schulze, R., Wilhelm, O., & Wittmann, W.W. (2000). Working memory capacity: Facets of a cognitive ability construct. *Personality and Individual Differences, 29*, 1017-1045.

Oberauer, K., & Süß, H.-M. (2000). Working memory and interference: A comment on Jenkins, Myerson, Hale, and Fry (1999). *Psychonomic Bulletin & Review, 7*, 727-733.

Oberauer, K. & Kliegl, R. (2001). Beyond resources – A model of complexity effects and age differences in working memory. *European Journal of Cognitive Psychology, 13*, 187-215.

Oberauer, K., Süß, R., Wilhelm, O., & Wittmann, W.W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence, 31*, 167-193.

Oberauer, K., Lang, E., & Engle, R.W. (2004). Working memory capacity and resistance to interference. *Journal of Memory and Language, 51*, 80-96.

Osaka, M., Nishizaki, Y., Komori, M., & Osaka, N. (2002). Effect of focus on working memory: Critical role of focus word in reading. *Memory & Cognition, 30*, 562-571.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition, 29*, 344-354.

Passolunghi, M.C. & Siegel, L.S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem-solving. *Journal of Experimental Child Psychology, 80*, 44-57.

Pearson, D., Logie, R.H., & Gilhooly, K. (1999). Verbal representations and spatial manipulation during mental synthesis. *European Journal of Cognitive Psychology, 11*, 295-341.

Potter, M.C., & Levy, E.I. (1986). Spatial enumeration without counting. *Child Development, 39*, 265-272.

Quinn, J., & McConnell, J. (1996). Irrelevant pictures in visual working memory. *Quarterly Journal of Experimental Psychology, 49A*, 200-215.

Ransdell, S., & Hecht, S. (2003). Time and resource limits on working memory: Cross-age consistency in counting span performance. *Journal of Experimental Child Psychology, 86*, 303-313.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55(A)*, 1339-1362.

Rao, S.C., Rainer, G., & Miller, E.K. (1997). Integration of the what and where in the primate prefrontal cortex. *Science, 276*, 821-824.

Reyna, V.F., & Brainerd, C.J. (1991). Fuzzy-trace theory and children's acquisition of mathematical and scientific concepts. *Learning and Individual Differences, 3*, 27-60.

Romani, C., & Martin, R.C. (1999). A deficit in the short-term retention of lexical-semantic information: Forgetting words but remembering a story. *Journal of Experimental Psychology: General, 128*, 56-77.

Roodenrys, S., Hulme, C.M., & Brown, G.D.A. (1993). The development of short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology, 56*, 431-442.

Saito, S. & Miyake, A. (2004). On the nature of forgetting and the processing-storage relationship in reading span performance. *Journal of Memory and Language, 50*, 425-443.

Salamé, 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 Behavior, 21*, 150-164.

Schneider, W., & Pressley, M. (1997). *Memory Development between two and twenty* (2nd ed.). Mahwah, NJ: Erlbaum.

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

Siegel, L.S. (1994). Working memory and reading: A life-span perspective. *International Journal of Behavioural Development, 17*, 190-124.

Slamecka, J.N., & Graf, P. (1978). The generation effect: Delineation of a phenomenon? *Journal of Experimental Psychology: Human Learning and Memory, 4*, 592-604.

Smyth, M., & Pendelton, L. (1989). Working memory for movements. *Quarterly Journal of Experimental Psychology, 41A*, 235-250.

Süß, H.-M., Oberauer, K., Wittmann, W.W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence, 30,* 261-288.

Swanson, H.L. (1999). What develops in working memory? A life span perspective. *Developmental Psychology, 35*, 986-100.

Swanson, H.L., Ashbaker, M.H., & Lee, C. (1996). Learning-disabled readers' working memory as a function of processing demands. *Journal of Experimental Child Psychology, 61*, 242-275.

Thorn, A. S. C., Gathercole, S. E., & Frankish, C. R. (in press). Redintegration and the benefits of long-term knowledge in verbal short-term memory: An evaluation of Schweikert's (1993) multinomial processing tree model. *Cognitive Psychology.*

Towse, J.N. & Hitch, G.J. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? *Quarterly Journal of Experimental Psychology, 48*A, 108-124.

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

Towse, J.N, Hitch, G.J., & Hutton, U. (2002). On the nature of the relationship between processing activity and item retention in children. *Journal of Experimental Child Psychology, 82*, 156-184.

Towse, J.N. & Houston-Price, C.M.T. (2001). Combining representations in working memory: A brief report. *British Journal of Developmental Psychology, 19*, 319-324.

Towse, J.N., Hitch, G.J., Hamilton, Z., Peacock, K., & Hutton, U.M.Z. (2005). Working memory period: The endurance of mental representations. *Quarterly Journal of Experimental Psychology A, 58*, 547-571.

Turley-Ames, K.J., & Whitfield, M.M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language, 49*, 446-468.

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

Vallar, G., & Baddeley, A.D. (1984). Fractionation of working memory. Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior, 23*, 151-161.

Vigliocco, G., Vinson, D. P., & Siri, S. (2005). Semantic similarity and grammatical class in naming actions. *Cognition, 94*, 91-100.

Wickens, D.D., Born, D.G., & Allen, C.K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 2*, 440-445.

Wickens, D.D., & Cammarata, S.A. (1986). Response class interference in STM. *Bulletin of the Psychonomic Society, 24*, 226-268.

Young, C.W., & Supa, M. (1941). Mnemic inhibition as a factor in the limitation of the memory span. *American Journal of Psychology, 54*, 546-552.

Appendix I

Stimulus items for Experiments 4 and 5

Processing items for Sentence Span task (Experiment 4)	Processing items for Operation Span tasks (Experiments 4 and 5)
Worms live in the ________	14 + 5 =
Shoes are worn on your ________	25 + 3 =
The moon shines at ________	12 + 3 =
Ducks swim on ________	31 + 8 =
A bicycle has two ________	15 + 4 =
Pigs have curly ________	33 + 2 =
A ship sails on the ________	16 + 2 =
A dog wags its ________	11 + 5 =
It gets dark at ________	24 + 5 =
A lift goes up and ________	31 + 3 =
A clock tells the ________	11 + 8 =
Giraffes have long ________	25 + 3 =
At bedtime I brush my ________	31 + 8 =
A chicken lays an ________	15 + 2 =
I use an umbrella when it ________	12 + 6 =
Gloves fit on your ________	33 + 2 =
Rabbits have long ________	16 + 2 =
Your teeth are in your ________	21 + 7 =
Santa comes down the ________	15 + 2 =
Nurses work in a ________	14 + 5 =
Sharks have sharp ________	11 + 5 =
David Beckham plays ________	25 + 3 =
A spider has eight ________	31 + 2 =
A mouse eats ________	11 + 8 =
A library has lots of ________	12 + 3 =
Aeroplanes fly in the ________	14 + 4 =
Wizards can cast ________	21 + 6 =

Storage items for Experiment 4	
5	Apple
8	Elephant
2	Baby
9	Banana
7	Teacher
2	Umbrella
5	Finger
3	Garden
4	Spiderman
7	Motorbike
3	Chicken
1	Holiday
5	November
7	Oranges
4	Flower
8	Caravan
4	Ice-cream
3	Lollipop
7	Aeroplane
2	Mother
5	Spider
3	Bicycle
5	Motorway
2	Table
7	Sister
4	Letter
8	Saturday

Processing items for Sentence Span (Experiment 5)	Storage items for Experiment5	
Oranges live in water	5	Horse
Roses smell nice	8	Green
Chairs lay eggs	2	School
Bananas have teeth	9	Foot
Shoes are worn on feet	7	Pipe
Apples grow on trees	2	Lake
Cars have wheels	5	Snow
Rabbits have long ears	3	Train
Bicycles eat grass	4	Ant
Elephants are big	7	Bear
Buses can talk	3	Rock
Dogs can bark	1	Mouth
Fish live in the ground	5	Blue
Ice-cream is hot	7	Car
Pianos play music	4	Belt
The sun is hot	8	Shoe
Bananas ride bicycles	4	Pond
Houses can sing	3	Rain
Your nose is on your face	7	Tail
Wheels are square	2	Box
Giraffes have long necks	5	Cup
Knives are soft	3	Cliff
Children go to school	5	Pink
Balls are round	2	Neck
Dogs can play guitar	7	Belt
Carrots are blue	4	Wind
Aeroplanes have wings	8	Dress

Appendix II

Stimulus items for Experiments 6-8

Words (taken from mean age-of-acquisition norms of Gilhooly and Logie, 1982)

bag
ball
band
bat
bay
beam
bed
bin
book
boot
box
dad
dart
dawn
deal
deck
deer
dip
dog
doll
door
dot
duck
dump
fat
fork
form
fox
frog
game
gang
gap
gas
girl
goal
gum
gun
hail
hall
hay
heap
hen
hill
home
hop
horn
jar
jaw
joke
joy
jump
lad
lamb
lamp
land
lap
law
lead
lift
lock
lord
luck
main
mast
men
mist
moan
mud
nail
net
nip
nod
page
palm
park
part
peck
peep
pest
pet
pig
pin
play
pond
pot
raid
rain
rake
rest
ring
road
rod
room
rope
rug
rust
salt
sand
shed
ship
soap
soil
sold
song
spot
sun
tack
tap
task
tear
tent
term
tin
tip
tool
toy
tuck
turn
van
walk
wand
west
win
wood
worm
yard

Nonwords (taken from the ARC Nonword Database of Rastle et al., 2002)

barg
bick
bleg
blit
blom
bock
bon
bool
bordge
borp
bup
dack
darp
deet
deg
derb
dern
dern
dirp
doob
dool
dop
dorge
dort
doz
ferg
fisp
fliss
fon
fusk
gab
ged
gell
gerp
gick
gol
goot
gos
hass
heb
hef
heg
hesk
hing
hish
hol
hom
jat
jeck
jisp
jit
jum
lafe
leck
leem
lerg
lib
lirm
lod
lolk
loog
lub
lud
mab
marn
mish
mord
mot
mun
nart
neeg
nerg
nug
pab
padge
pag
parn
peeb
peem
pib
pim
pleg
pock
pook
poy
pud
rab
reb
reeb
ref
rerb
resk
rilk
rop
rorm
rosh
rost
sarc
sarm
sarp
sep
snoy
sodge
speep
stib
susk
tam
tarb
tep
terch
tob
toock
tord
torm
toz
tudge
tunk
turg
vont
wast
weff
wek
wirp
woodge
worl
yoam

Stimuli with onset phoneme /k/

Words	Nonwords
cage	kafe
camp	kam
can	kark
cap	kav
car	keb
cart	ked
case	kef
clay	keem
coal	kib
coat	kig
cod	klat
cone	koll
cup	kom
cut	koob
keen	koom
kerb	korm
kick	kud
kilt	kug