THE UNIVERSITY OF
WARWICK

Age-Related Changes in Associative Memory

By

Stephen P. Badham

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Psychology

University of Warwick, Department of Psychology
September 2011

Dedication

To my loving wife Stephanie

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my first supervisor Elizabeth Maylor for providing valuable advice and guidance throughout my PhD. I would also like to thank my second supervisor Zachary Estes for providing advice and guidance on Chapters 6 and 9 and to acknowledge Neil Stewart and Kimberly Wade for providing advice on Chapter 4 and Moshe Naveh-Benjamin for providing advice on Chapter 5. I wish to thank Lauren Brawn and Charlotte Gillingham for collecting some of the data reported in Chapter 3, Hannah Watts and Natalie Woods for collecting some of the data reported in Chapter 4, Laura Steel and Katherine Tyler for collecting some of data reported in Chapters 5 and 8, and Charlotte Gallimore and Jennifer Harber for collecting the data reported in Chapter 7. I also wish to thank Tamra Bireta for providing data analysed in Chapter 3, Theodora Zarkadi for providing stimuli used in Chapter 4, Moshe Naveh-Benjamin for providing some of the stimuli used in Chapters 5 and 8, and Myra Fernandes for providing some of the stimuli used in Chapter 8. Finally, I would like to thank all of the participants who volunteered to take part in my research.

Declaration

I hereby confirm that I completed this thesis independently, that I have not heretofore presented this thesis to another department or university, and that I have listed all references used, and have given credit to all additional sources of assistance.

Note on Inclusion of Published Work

Experiments 4, 5 and 6 of this thesis have been published during the period of my PhD registration, and the copyright of these papers resides with the publishers (the reproduction of the papers in this thesis is permitted under the terms of the copyright agreement). The publications are:

Badham, S. P., Estes, Z., & Maylor, E. A. (2011, May 30). Integrative and semantic relations equally alleviate age-related associative memory deficits. *Psychology and Aging*. Advance online publication. doi: 10.1037/a0023924.

Badham, S. P., & Maylor, E. A. (2011). Age-related associative deficits are absent with nonwords. *Psychology and Aging, 26*, 689-694.

Abstract

Older adults suffer from many cognitive impairments relative to young adults and one of the most established types of age-related cognitive decline is a reduction in memory performance. Memory for single units of information (item memory) have been shown to be less susceptible to cognitive ageing than memory for associations among units of information (associative memory). An associative deficit hypothesis has been used to describe these observations as an age-related impairment in forming links between single units of information. The thesis elucidated specific differences between item and associative memory and evaluated how such differences correspond to their differential susceptibility to the effects of cognitive ageing. This indicated links between the associative deficit hypothesis and other theories of age-related memory decline, in particular, to the notion of age deficits in memory resulting from age deficits in self-initiated processing (in the absence of environmental support).

Experiments 1-3 considered associative memory where the processing of associations was encouraged by distinctiveness of memory stimuli. Environmental support provided by distinctiveness was shown to improve associative memory in older adults. Experiments 4-7 considered how item and associative memory differ in their support from preexisting knowledge. Experimentally equating preexisting knowledge for item and associative memory tests eliminated the age-related associative deficit. Furthermore, it was found that preexisting knowledge could be used to enhance associative memory performance in older adults by providing support to encoding and/or retrieval processes. Experiment 8 established that item and associative memory processes were equally disrupted by a concurrent task, which indicated that both memory types are similarly affected by levels of available cognitive resources. In general, age-related associative deficits were considered to result from differing levels of environmental support for item and associative memory as opposed to a differential decline of item and associative memory processes.

Chapter 1: An Overview of Cognitive Ageing

This chapter explores cognitive ageing research in general and provides a background for research reported in the thesis. Existing theoretical accounts of cognitive ageing will be summarised in order to clarify how the current research into age-related associative deficits sits within the wider literature.

Over the last century, life expectancy has continued to increase with advances in quality of living due to changes in factors such as income, nutrition, sanitation and medication (Riley, 2001). In the early 20th century, rises in life expectancy were mainly due to reductions in infant and child mortality. However, from the mid 20th century onwards, increases in life expectancy were largely due to improvements in survival over age 65 (Oeppen & Vaupel, 2002). This has led to a large increase in the number of people reaching old age and coincides with an increasing need for geriatric research. In particular, the quantity of cognitive ageing research has increased rapidly in recent years (Salthouse, 2010) as there is a greater requirement to understand and combat cognitive decline in old age.

Stereotypical views of cognitive ageing regard older adults to have poorer memories than young adults and to be generally less competent at cognitive tasks than young adults (Hertzog & Hultsch, 2000). Unsurprisingly, such views are supported by empirical cognitive research and age differences in memory and cognitive resources provide a key focus for theorists (Zacks, Hasher, & Li, 2000). Healthy older adults show deficits in a range of cognitive tasks compared to young adults (see Park, 2000; Salthouse, 2010; Verhaeghen, Marcoen, & Goossens, 1993, for reviews). Older adults are slower than young adults across various measures of speed of cognition (e.g., Salthouse, 1996), they have a reduced working memory capacity compared to young adults (e.g., Craik, 2000; Craik & Byrd, 1982), and they show reduced ability to focus attention compared to young adults (e.g., Hasher & Zacks, 1988). Another area where older adults show deficits relative to young adults is in the formation of episodic memories (see Spencer & Raz, 1995, for a review). More recently this has been hypothesised to be a result of a specific age-related deficit in associating items in memory (Naveh-Benjamin, 2000). Unlike other cognitive factors that decline with old age, age-related associative deficits are yet to be fully reconciled with the notion of cognitive decline resulting from a global decline in processing resources. This is because associative deficits are specific to one type of memory, and are not replicated in young adults under divided attention[1] (e.g., Naveh-Benjamin, Guez, & Shulman, 2004). The research reported in this thesis aimed to clarify and test the nature of associative memory, and to establish whether associative memory deficits are truly dissociated from global memory deficits.

In general, age deficits are hypothesised to be due to global age-related decline in cognitive functioning. In line with this view, Craik and colleagues (e.g., Craik, 2000; Craik & Byrd, 1982) developed a resource deficit hypothesis whereby age-related decline is at its most extreme for tasks where cognitive demands are highest. In relation to memory tasks, Craik (1986) hypothesised that increasing environmental support reduces cognitive demands and therefore reduces age deficits, whereas increasing self-initiated processing increases cognitive demand and therefore increases age deficits. Table 1 shows how age deficits in memory were hypothesised by Craik (1986) to correspond to environmental support and self-initiated processing.

---

[1] There is also some evidence for the opposite, where divided attention does produce associative deficits in young adults. See Chapter 8 for a full review of associative deficits in young adults under divided attention.

Table 1

Environmental Support, Self-Initiated Activity, and Age Deficits for Different Types of Memory Task

Task	Environmental support	Self-initiated activity	Age-deficit
Remembering to remember	lower	higher	higher
Free recall	↓	↑	↑
Cued recall			
Recognition			
Relearning			
Procedural memory	higher	lower	lower

Note. Adapted from "A Functional Account of Age Differences in Memory," by F. I. M. Craik, 1986, In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities, mechanisms, and performance,* p. 412, Amsterdam. Copyright 1986 by Elsevier Science Publishers.

A decline in cognitive resources across the lifespan is perhaps the most appealing explanation for observed differences between young and older adults. This is because it is a single approach, which can be used to explain a wide range of data and is consistent with the notion of decline where older adults are typically observed to be poorer at cognitive tasks than young adults. One of the reasons that reductions in cognitive resources across the lifespan can explain a wide variety of data is that they can be interpreted in different ways. This has given rise to several theoretical mechanisms of age differences in cognition, which have been used to explain age differences in performance between young and older adults in cognitive tasks. The most popular accounts are summarised in the rest of this chapter.

Processing Speed

The processing speed theory of cognitive ageing developed by Salthouse (1991, 1996), which builds on earlier work by Birren (1965), argues that older adults perform more poorly than young adults on a range of cognitive tasks because they are slower at processing information. Processing speed is typically measured by perceptual speed tasks such as the Digit Symbol Substitution Task (Wechsler, 1981); it has been widely demonstrated that older adults are slower than young adults using measures of perceptual speed where participants are required to make rapid judgements about the similarity or differences between symbols or digit/letter strings (Salthouse, 1996). In his article, Salthouse (1996) presented a large range of evidence to suggest that cognitive performance across a wide range of tasks can be explained by the rate at which individuals process information. This is consistent with an earlier review by Cerella (1985) which also showed reliable age-related cognitive slowing across a range of studies. Therefore this evidence has led cognitive slowing to become a dominant theory within the literature to explain patterns of cognitive ageing (Fisher, Duffy, & Katsikopoulos, 2000).

The theory has two major components - the limited time mechanism and the simultaneity mechanism. The first component postulates that if processing is slower, then there is less time to perform cognitive operations. In its simplest form, this mechanism indicates that older adults will require more time than young adults to complete a given task. This is evident in tasks where difficulty is low and individual differences are measured by speed of completion (Salthouse, 1996). For more complex tasks, where performance is measured by accuracy or memory output, where the *quality* of processing is important, the mechanism can be viewed slightly differently. Such tasks may require a variety of processes such as associations, elaborations and rehearsals, and with slower cognitive performance fewer of these operations may be able to take place before output is required. In particular, with regards to memory, slower cognition may reduce the number of rehearsals an individual can make before memory retrieval. This means that if one process is dependent on the output of another, with limited time, the quality of earlier output may be restricted and affect overall task performance. Age differences are larger for more complex tasks requiring multiple processes (e.g., Naveh-Benjamin, 2000; Salthouse, 1991) and this mechanism may provide a description of this effect. Ultimately, more processing will often result in better performance and the amount of processing completed will depend on processing speed (Salthouse, 1996).

The second component, the simultaneity mechanism, postulates that the products of earlier processes may be lost by the time that later processing is completed. This mechanism assumes that information degrades over time, which can be manifested as a loss of quality or quantity of that information - memory fades over time (e.g., Peterson & Peterson, 1959) and theoretically so will the products of cognitive processes. This means that the slowing of cognitive processes (e.g., elaboration, search, rehearsal, retrieval etc) will cause information to be accessed at a later time, a time when earlier information may be lost or degraded. The fact that the mechanism describes a disruption of available information could also apply to the disruption of working memory; speed of processing has been demonstrated to account for a large proportion of individual differences in working memory (Salthouse, 1996). Disruption of working memory capability in old age is another theoretical account of age differences in cognition that is explored in more detail in the next section of this chapter. Notably, the simultaneity mechanism is based on internal limitations; therefore, externally increasing the time available for task completion with older adults will not necessarily lead to improved performance. This means that effects of cognitive slowing can be viewed as global in that they are not limited to tasks where speed is a dependent variable.

Cognitive slowing also has a physiological basis - it has been shown that white matter deterioration in old age correlates with cognitive slowing (see Gunning-Dixon & Raz, 2000, for a review). White matter deterioration (the formation of white matter hyperintensities) occurs in the ageing brain, possibly arising from vascular and neural pathologies (Dennis & Cabeza, 2008). The deterioration may result in a reduction of neurotransmission speed in older adults, which is observed behaviourally as age-related slowing in cognitive tasks (Gunning-Dixon & Raz, 2000).

Overall, processing speed has been shown to explain a large proportion of age-related variance across a range of cognitive tasks such as working memory, free recall, spatial abilities and reasoning (e.g., Salthouse, 1993). Older adults are also reliably found to be slower than young adults for different measures of speed based on perceptual classification tasks. As processing speed is an approach that is strongly supported by the literature, the majority of experiments reported in this thesis contain a measure of processing speed (the Digit Symbol Substitution Task, Wechsler, 1981) to clarify individual differences in performance.

Working Memory

An age-related impairment of working memory ability can also be used to explain a variety of experimental observations. Craik and colleagues have proposed that older adults show a decline in attentional resources leading to reduced working memory capacity (e.g., Craik, 2000; Craik & Byrd, 1982) and this view is considered as highly influential in ageing research (Zacks et al., 2000). Working memory capacity is defined as the amount of resources available at a given moment in time to engage in mental operations that manipulate information in short-term memory, that is, storage, retrieval and transformation of information (Baddeley, 2003).

Age differences are relatively small in simple short-term memory tasks requiring the temporary retention of information in memory; however, when short-term memory tasks require manipulation of stored material, or alternation between processing different information, age differences become much greater (Craik, 2000). These observations provide evidence that distinguishes between short-term memory deficits and working memory deficits. More precisely, Craik and Byrd (1982) proposed that older adults are deficient at self-initiated processing and that age differences in working memory can be alleviated when there is sufficient environmental support. This is evidenced by studies that show smaller age differences when external cues can be used to support task performance. The notion of environmental support can explain patterns of age differences in tasks with different memory retrieval measures: Age differences are typically larger (with older adults performing more poorly) for recall memory tests than recognition memory tests (e.g., Craik & McDowd, 1987; Light, Prull, La Voie, & Healy, 2000; Schonfield & Robertson, 1966). With recognition memory tests, information that was present during encoding is also present during retrieval. This provides environmental support to the retrieval process, which supposedly helps older adults to improve their memory performance relative to young adults (Craik, 1986). Also, when questionnaires are presented auditorily, age differences are more apparent than when they are presented in a written format. This is most likely because the information about the question and response options have to be retained in memory for auditory questionnaires as it is not readily available as with written questionnaires (Park, 2000). The notion of increased environmental support reducing age differences in cognitive tasks can also be applied to cued recall. When the cue is semantically related to the target, age differences are smaller than when the cue and target are unrelated (e.g., Naveh-Benjamin, 2000). This is theoretically because semantic relations between cues and targets provide extra environmental support that benefits older adults' memory performance more than young adults'. In general, memory cues at encoding and prompts at retrieval provide successful support to older adults' memory (Park, 2000).

There are different mechanisms hypothesised for production of age-related deficits in working memory. Firstly, the most direct hypothesis is that the fundamental capacity of working memory is reduced in conjunction with an age-related decline in attentional resources (Craik & Byrd, 1982). Secondly, as discussed earlier, speed of processing may have an impact on working memory by reducing the number of operations that can be completed before information in working memory fades over time (Salthouse, 1996). Thirdly, a reduced ability in older adults to inhibit irrelevant information (which is discussed below) may mean that working memory becomes cluttered with irrelevant information, reducing the space available for task-relevant information (Hasher & Zacks, 1988). The slowing and inhibition views produce an indirect reduction in working memory mediated by more fundamental age deficits but the capacity view is a direct reduction in working memory due to age-related cognitive decline. This capacity based view, which is distinct from the other accounts, therefore provides a separate approach to understanding cognitive ageing. However, the majority of evidence supporting a reduction in working memory with increased age does not distinguish between the underlying mechanisms responsible for that reduction.

A potential issue with the working memory views of cognitive ageing is that they do not link tightly with the three-component model proposed by Baddeley and Hitch (1974), which is currently the dominant theory of working memory (Barrouillet, Portrat, & Camos, 2011). It consists of a modality-free central executive and two subordinate modality-specific components (the visuo-spatial sketch pad and the articulatory loop). Despite its effectiveness at explaining a variety of patterns of working memory performance, few links have been made between the three-component model and working memory decline in old age (Zacks et al., 2000). The central executive has been considered in relation to cognitive ageing, largely because it is considered to be responsible for executive functioning, which does show decline in old age (Parkin & Java, 2000; West, 1996). Physiologically, executive functioning is widely considered to be linked to prefrontal cortex functionality (e.g., Raz, 2000). Prefrontal degradation as a result of healthy ageing has been found to occur earlier than degradation of other areas of the brain (e.g., Dennis & Cabeza, 2008; West, 1996). Therefore the health of the prefrontal cortex and its impact on executive functioning and working memory are considered as important factors in cognitive ageing research. In addition to working memory, executive functioning and its mediation via the prefrontal cortex can also be applied to inhibition theories (West, 1996), which are reviewed below.

Inhibition

Another major theory in cognitive ageing sees age deficits in many tasks as deficits in inhibitory functionality (Hasher & Zacks, 1988). The key idea is that older adults are less able to focus attention on relevant material (i.e., less able to inhibit irrelevant information) and that their attentional resources are diffused compared to young adults during cognitive tasks. Failure to inhibit irrelevant behaviour is evident in older adults anecdotally as they are often considered to ‘speak their minds’ and make occasional inappropriate remarks (Park, 2000). There is also a body of empirical evidence to support the theory: Experimental evidence includes the absence of negative priming in older adults (e.g., Hasher, Stoltzfus, Zacks, & Rypma, 1991; McDowd & Oseas-Kreger, 1991). This is where information that should be inhibited in relation to an earlier trial but then becomes relevant in a later trial causes slowed access to that information in the later trial: Young adults showed slowed access but older adults did not, indicating that information was not sufficiently inhibited by older adults in the earlier trial.[2] There are studies that show an increase in the fan effect with age (Cohen, 1990; Gerard, Zacks, Hasher, & Radvansky, 1991). The fan effect occurs when the more associations there are to a concept, the slower and more error prone access to that concept is. Age increases in the fan effect have been attributed to inhibitory deficits in old age (Gerard et al., 1991; Zacks et al., 2000). This is because the fan effect is assumed to be driven by interference at retrieval, interference that is hypothetically greater in older adults due to inhibitory deficits. Older adults are found to be more susceptible than young adults to proactive interference, where irrelevant information from earlier trials reduces performance during later trials (e.g., Lustig, May, & Hasher, 2001). There is also evidence to show that older adults are poorer than young adults at directed forgetting, which indicates that they are less able to consciously suppress items in memory (Zacks, Radvansky, & Hasher, 1996).

Hasher and Zacks (1988) identified three mechanisms by which inhibitory deficits can impact cognitive functioning: *access*, *deletion* and *restraint*. These mechanisms are all factors that control the contents of working memory (Yoon, May,

[2] However, some studies have shown equivalent negative priming in young and older adults (e.g., Kramer, Humphrey, Larish, Logan, & Strayer, 1994; see May, Kane, & Hasher, 1995; McDowd & Shaw, 2000, for reviews).

& Hasher, 2000). Firstly, inhibitory mechanisms are assumed to be necessary to prevent irrelevant information from entering working memory, only allowing *access* to relevant information. Secondly, inhibition is also assumed to be responsible for *deletion* of irrelevant information that is already in working memory but that is no longer relevant to ongoing tasks. These two mechanisms together are used to ensure that only information appropriate to current goals is present in working memory (Zacks et al., 2000). These factors elucidate the point made earlier that older adults' working memory capacity may be reduced by inhibitory deficits. This is because for older adults more uninhibited, irrelevant information may enter and stay in working memory, leaving less space for relevant information. Along these lines, the evidence for reduced working memory capacity in old age may provide circumstantial support for the inhibition deficit hypothesis. Finally, inhibitory mechanisms are assumed to be necessary for *restraint*, which is considered responsible for preventing dominant responses from being activated before they are fully evaluated against less probable but more appropriate responses (Yoon et al., 2000). This restraint mechanism has been argued to produce the most pronounced age differences in cognition relevant to inhibitory deficits (Zacks & Hasher, 1997).

The reduced inhibition theory has gained some opposition since its original introduction, where the global nature of an inhibition deficit could not explain preserved inhibition for certain tasks (e.g., Burke, 1997; McDowd, 1997). However it remains to be discredited and provides a more specific framework by which to view reductions in working memory capacity. In addition, it is reflected physiologically by the presence of inhibitory neural circuits, which adds to the plausibility of the theory (Park, 2000). One aspect of cognitive ageing that appears to go against inhibitory theory is a reduction in context memory in older adults compared to young adults above and beyond age reductions in content memory (Light, 1991; Spencer & Raz, 1995). This is particularly relevant to the present thesis because a reduction in context memory with age can be explained by age-related associative deficits (this is reviewed in the following chapter). Older adults often remember less about context than young adults, even when context information is irrelevant (e.g., Naveh-Benjamin 2000), therefore indicating that irrelevant information was forgotten more in older adults than in young adults. If older adults had inhibition deficits, one would expect them to remember *more* irrelevant information.

Sensory Function

There has been evidence to suggest that age-related declines in sensory function may be responsible for age-related declines in cognition (Schneider & Pichora-Fuller, 2000). In simple terms, if participants cannot accurately perceive test stimuli, then their performance will suffer. In reality (and in the current thesis), experimenters will usually aim to make stimuli salient enough so that participants of differing sensory abilities can detect them (Schneider & Pichora-Fuller, 2000). Despite this, however, older participants may expend greater effort in perceiving stimuli and this could affect their cognitive performance. For example, Rabbitt (1968) showed that young adults were less able to later recall digits heard with white noise in the background compared to digits heard clearly, despite the fact that they were able to accurately repeat (i.e., perceive) all of the digits as they were encountered.

Lindenberger and Baltes (1994) tested a range of older adults with a battery of tests covering five different cognitive areas (speed, reasoning, memory, knowledge, and fluency). They also measured visual and auditory acuity and found that nearly all of the age-related variance in the cognitive tests could be accounted for by measures of sensory functioning. A particularly surprising result was that measures of sensory functioning were as successful as speed at predicting cognitive ability. Rather than concluding that sensory functioning directly affects cognition, they hypothesised that sensory and cognitive functioning were affected by a common underlying factor. A further study by Lindenberger and Baltes (1997) also found a strong relation between sensory and cognitive functioning in older adults. In addition, sociobiographical indicators such as education and social class did not predict cognitive ability as strongly as sensory functioning. This indicated that biological factors provide a stronger measure of cognitive integrity than social history, adding weight to the importance of the relationship between cognition and sensory functioning. Salthouse (2010) also found that age-related degradation of sensory functioning could account for declines in cognitive ability (speed, fluid intelligence, memory) across age.

There are three key views as to how sensory functioning and cognition may interact (Schneider & Pichora-Fuller, 2000). The first is the common cause view, indicated above, where both sensory and cognitive functioning decline as a result of some underlying neural degradation. The second is the multiple causes view, where both cognitive and sensory systems become degraded by age for separate reasons, which happen to correlate when measured. A third view is that of perceptual degradation, where reductions in sensory function lead to cognitive decline. This third view is in line with neuroimaging data that suggest that cognitive resources may be used to compensate for sensory degradation. Activity in the prefrontal cortex has been shown to increase with age to a similar extent that visual cortex activity has been shown to decrease with age (Dennis & Cabeza, 2008). This could mean that older adults are using top-down processes to make sense of their surroundings in order to compensate for a reduction in sensory function. Rabbitt (1991) also hypothesised that increased effort needed to perceive less salient stimuli reduced participants' ability to rehearse or elaborate memory stimuli during encoding. Although speculative at this stage, it may be the case that decline in sensory function can result in cognitive slowing as older adults take more time to make sense of their surroundings – leading to cognitive difficulties outlined above attributed to general slowing. Additionally, or alternatively, older adults may be using working memory capacity to compensate for sensory deficits, which could explain the poorer working memory performance observed in older adults outlined above. Ultimately, sensory function is clearly poorer in older adults compared to young adults. However, it remains to be firmly established which factors cause this degradation and how it impacts on general cognitive ageing.

Dual Process Accounts of Memory

Dual process theory is the final major theory relevant to cognitive ageing that will be discussed before directly addressing age-related associative deficits in the following chapter. The theory describes how age differences differ between familiarity and recollection measures. In dual process models, familiarity and recollection are considered as different memory processes (Yonelinas, 2002). Familiarity is seen as more automatic, providing a sense that a given stimulus has been encountered before following a recognition probe, whereas recollection is seen as controlled conscious retrieval of specific episodic experiences. Dual process models commonly assume that during recognition memory tasks, when the level of familiarity/unfamiliarity in memory is ambiguous, a further recollection based memory search is required before a response can be made (Yonelinas, 2002). These two processes have been shown to be differentially affected by age; tasks involving mainly familiarity typically show minimal age deficits whereas tasks involving recollection typically show more pronounced age deficits (Light et al., 2000).

The process dissociation procedure (Jacoby, 1991) has been used to estimate young and older adults' levels of familiarity and recollection. In this procedure, participants study items and are tested in different ways to establish levels of familiarity and recollection performance. Items to be memorised differ in their method of presentation so as to form two groups (e.g., half of the items may be presented visually and half auditorily, or half of the items may be presented in one list and half in another). Participants then complete two recognition tests. One test is a standard inclusion recognition test where participants are asked to respond positively to items they have studied before and negatively to new items. Positive responses on this test are presumed to be based on both familiarity and recollection of the old items. The other test is an exclusion recognition test, where participants must respond positively only if the item was presented in a certain way (e.g., respond positively only to items previously heard but not to items previously seen). Erroneous responses on this test are assumed to be entirely due to familiarity as participants have falsely endorsed items because they could not recollect their original method of presentation. This provides a measure of familiarity, and the measure of recollection is obtained by subtracting this familiarity from positive responses in the first test (which are due to both recollection and familiarity). Such methods have yielded larger age differences in recollection than familiarity (e.g., Benjamin & Craik, 2001; Jacoby, 1999; Jennings & Jacoby, 1993).

A second method to measure familiarity and recollection separately is to ask participants directly via the remember/know procedure developed by Tulving (1985).

In brief, participants study some memory stimuli and then complete a recognition test; when they respond positively to a seen-before item, they are asked to make a remember/know judgement about the basis of their decision. Participants are asked to indicate a *remember* judgement if they specifically recall the episodic event of encountering the item in the original memory set. They are also asked to indicate a *know* judgement if they do not necessarily remember encountering the item in the original memory set but sense that it feels familiar enough for them to decide that they must have seen it earlier. Thus, these judgements allow experimenters to categorise responses based on recollection (remember) and familiarity (know). The pattern of age differences is the same as that found with the process dissociation procedure – larger age differences for recollection than familiarity (see Light et al., 2000; Yonelinas, 2002, for reviews).

The source of larger age differences in recollection than familiarity is not entirely clear. There is considerable evidence to suggest that older adults are less likely than young adults to implement encoding strategies (e.g., Luszcz, Roberts, & Mattiske, 1990; Witte, Freund, & Sebby, 1990), and encouraging the implementation of encoding strategies has been shown to attenuate age-related memory deficits (Naveh-Benjamin, Brav, & Levy, 2007; Park, Smith, Morrell, Puglisi, & Dudley, 1990; Treat & Reese, 1976). This indicates that older adults may be performing poorly at recollection because they are not spontaneously using any strategies to aid recollective processes. Other research has demonstrated that less effortful processes (i.e., familiarity) show smaller age-related decline (e.g., Hasher & Zacks, 1979; Salthouse, 1988). This is in line with the working memory and inhibitory deficit theories outlined above, where older adults show a reduced capacity for control of attention. In addition, a range of dual process models of memory agree that familiarity is faster than recollection (Yonelinas, 2002) so age-related recollection deficits may stem from cognitive slowing.

An issue with the conclusion that there is a differential effect of age between recollection and familiarity is that methods of measuring the variables may not be pure. It could be that recollection and familiarity work together and not independently and this would undermine the dual process account (Light et al., 2000). There are also some studies that demonstrate no differential age effect between recollection and familiarity when estimates of recollection are high, although this pattern of results has been attributed to ceiling effects (Yonelinas, 2002). Finally, and most importantly with respect to this thesis, dual process accounts of age differences in memory are not clearly distinguishable from age-related associative deficits. That is, measures of familiarity are similar to measures of item memory (recognising a single stimulus) and measures of recollection are similar to associative/context memory (remembering the source/context that the stimulus was encountered). Therefore dual process accounts of age differences in memory may be explained by age-related deficits in forming associative memories (i.e., age deficits in binding units of information to the context in which they were originally encountered).

Summary

This chapter has explored five major theories of cognitive ageing that have been applied to age-related deficits in episodic memory; it has also introduced some of the approaches and ideas that will later be used to understand age-related associative deficits. The associative deficit hypothesis (Naveh-Benjamin, 2000) is also a major theory of cognitive ageing that provides a different theoretical approach to cognitive ageing research. This is the main approach discussed throughout the thesis and will be reviewed in detail in the following chapter.

It has been shown that the different accounts of memory deficits in later life are heavily inter-related and not necessarily distinct. Recollection deficits may be a result of working memory or inhibitory deficits. Working memory deficits may stem from inhibitory deficits or cognitive slowing. Cognitive slowing may stem from sensory deficits, and working memory may compensate for sensory deficits. Memory research related to cognitive ageing still has a long way to go and the general approach in the literature is to clarify and unify existing theories. This ethos is continued here and the aim of this thesis is to clarify and understand associative deficits to see where they fit within the current understanding of age-related memory decline.

Chapter 2: Age-Related Associative Deficits

The key age-related cognitive change that will be considered throughout this thesis is the ability to form associations between units in memory. This ability is particularly susceptible to the ageing process and older adults show reductions in associative memory ability relative to young adults (Chalfonte & Johnson, 1996; Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008a). The focus of the thesis is to clarify and elaborate our understanding of associative memory and to test what factors mediate age differences in associative memory ability. The current chapter describes existing research into age-related associative deficits and provides a summary of findings and hypotheses related to research conducted throughout the thesis.

Age-related associative deficits have been found in early research (e.g., Gilbert, 1941) and more recent studies (e.g., Dunlosky & Hertzog, 1998; Naveh-Benjamin, 2000, Exp. 4) using cued recall of word pairs. Participants were shown pairs of words for memorisation in a study period. Following this, they completed cued recall tests where they were shown one word of each pair and were asked to recall the other. Young adults performed better than older adults in these tests, indicating that older adults struggle to form associative links between stimuli that they encounter. A problem with the interpretation of these results is that they do not distinguish between associative deficits and general memory deficits in older adults relative to young adults. Underwood (1969) hypothesised that an episode of memory contains associative links between various attributes (e.g., temporal, spatial, contextual); encoding and retrieving the episode requires knowledge of the attributes and their relations to each other. This is a dominant view of episodic memory in the literature (Naveh-Benjamin, 2006) and provides a distinction between memory for associations connecting units of information and memory for the units individually. Many previous studies have also argued for this distinction (e.g., Chalfonte & Johnson, 1996; Humphreys, 1976; Trinkler, King, Spiers, & Burgess, 2006).

A meta analysis by Spencer and Raz (1995) addressed memory for context and content information in young and older adults. It was the first study to thoroughly review memory for units of information compared to memory for association among units of information between young and older adults. The analysis considered 46 studies and the primary result was clear: Age differences were larger for context memory than for content memory. Content and context memory stimuli ranged across the studies analysed. Examples of content memory were memory for words, actions and objects. Examples of context memory were temporal positions, spatial locations, modality and colour. Spencer and Raz (1995) explored a range of ideas to explain the pattern of results. They considered the possibility that contextual information was less goal-relevant and received less attention than content memory - to the greater detriment of older adults' memory relative to young adults' memory. It was hypothesised that this was mediated by difficulties that older adults have with focus of attention, working memory capacity and inhibition, which may have prevented them from focusing on/applying resources to contextual information. They argued that memory for context may be mediated by the prefrontal cortex to a much greater extent than memory for content, highlighting its role in temporal order memory and working memory. They also considered the role of metamemory such that as older adults adapt to poorer overall memory ability, they tend to use more reliable, less elaborate encoding and retrieval strategies, which are less favourable to contextual information than content information.

Following this research, Naveh-Benjamin (2000) more specifically considered age-related associative deficits and formed the associative deficit hypothesis (ADH). Using a paradigm similar to that of Humphreys (1976), Naveh-Benjamin separately tested memory for units of information (items) and associations between those units in young and older adults. The ADH specified that older adults have particular deficits at forming associations between items of memory. In a single paper (using within-subjects measurements of item and associative memory), Naveh-Benjamin compared the magnitude of age deficits for item and associative memory and found that associative memory age deficits were significantly larger than item memory age deficits.

Naveh-Benjamin's (2000) Experiment 1 tested memory for items (words and nonwords) and associations (word-nonword pairs). Participants were shown word-nonword pairs sequentially and then after a short delay were tested on their memory for words, nonwords and word-nonword associations in three different recognition tests. For the words recognition test, participants were shown words and had to indicate if they had seen them before in the original memory set. Half of the test words were old and half were new (not presented before). A similar test was conducted with old and new nonwords. To test word-nonword pairs via recognition, participants were again presented with word-nonword pairs. Half were old and appeared exactly as they had during the study period and half were recombined including a word and a nonword that were originally presented in different pairs. This meant that participants could not make an old/new judgement for associative memory on the basis of familiarity with individual words and nonwords as both old and recombined pairs used seen-before components. The experiment therefore had a recognition test purely based on associative memory with which to compare to recognition tests purely based on item memory. In general, older adults performed poorer than young adults at all tests. However, age differences were not significant for word memory but were significant for nonword memory and word-nonword associative memory. This resulted in an interaction between memory test and age where older adults showed greater associative deficits than item deficits (item deficits based on word memory) in comparison with young adults.

Naveh-Benjamin's (2000) Experiment 2 avoided the complication caused by using nonwords as memory stimuli by presenting pairs of words at study. Again, using recognition tests of item (word) and associative (word-word) memory, older adults showed significantly greater memory deficits compared to young adults in associative memory relative to item memory. (Experiment 3 confirmed the findings with different stimuli using similar item and associative recognition tests.) Experiment 2 also manipulated study instructions (incidental vs. intentional learning). When participants were instructed to memorise associations, the age-related associative deficits were more pronounced than when participants were instructed to just focus on individual words (i.e., when they were expecting an item test only). This indicated that older adults were less able to apply a strategy than young adults: Intentional encoding of associations allowed young adults to increase their associative memory performance more than older adults when compared to incidental learning of associations. Therefore, this indicated that young adults were consciously incorporating some sort of strategy to improve associative memory more successfully than were older adults. (Older adults' associative memory was still worse than young adults under both incidental and intentional learning of associations.) Naveh-Benjamin hypothesised that the effect may be based on both prefrontal (strategic) deficits and hippocampal (binding) deficits in older adults relative to young adults. Strategic deficits in older adults were also apparent when participants were questioned post test. Young participants reported the use of sentence production to link words at study whereas older participants tended not to use any strategy and older participants who did use a strategy tended to rely on a basic rehearsal (repetition) strategy.

Following Naveh-Benjamin's (2000) ADH, many studies can be found that have measured both item and associative memory in young and older adults. Older adults have shown larger deficits in associative than item memory relative to young adults with a range of stimuli including associations between word pairs (e.g., Castel & Craik, 2003; Light, Patterson, Chung, & Healy, 2004; Naveh-Benjamin et al., 2007; Naveh-Benjamin, Guez, & Shulman, 2004), words and fonts (Naveh-Benjamin, 2000, Exp. 3), pairs of pictures (Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003, Exp. 1), objects and locations (Mitchell, Johnson, Raye, Mather, & D'Esposito, 2000, see also Chalfonte & Johnson, 1996), pairs of faces (Bastin & Van der Linden, 2006), faces and spatial locations (Bastin & Van der Linden, 2006), faces and temporal presentation (Bastin & Van der Linden, 2005), and faces and names (e.g., James, Fogler, & Tauber, 2008; Naveh-Benjamin, Guez, Kilb, & Reedy, 2004).

Another line of evidence suggesting that age differences in associative memory are dissociated from general age differences in memory comes from dual task experiments. In the previous chapter it was discussed how reduced cognitive resources in old age are theorised to produce many of the observed patterns of cognitive aging. Dual task experiments are designed to tax cognitive resources and many age-related cognitive deficits can be increased under dual task conditions (Kramer & Madden, 2008). This demonstrates that dual task experiments tax cognitive processes where older adults already express deficits. Young adults' memory can also be reduced experimentally by requiring them to complete a concurrent task whilst encoding and/or retrieving information. Contrary to expectations, when young adults' memory is reduced in this way, their item and associative memory is usually hindered to the same extent by a concurrent task.

Naveh-Benjamin, Guez, and Shulman (2004) measured item and associative recognition memory based on pairs of words. Young and older adults were tested and the young group completed the experiment under full attention and under divided attention during encoding. When older adults were compared to young adults under full attention, their overall memory was lower and their associative memory was disproportionately lower than their item memory (i.e., support for an ADH). When older adults were compared to the young adults under divided attention, their overall memory was comparable, but again their associative memory was lower than their item memory, resulting in a significant interaction between age and memory test (item/associative). This shows that reducing memory performance by dividing attention in young adults did not result in them showing associative deficits like those observed in older adults. Many studies have shown similar results but other studies have found that dividing attention does impact associative memory more than item memory. In order to clarify these discrepancies, Chapter 8 explores the role of divided attention in young adults' item and associative memory with a different concurrent task to that used in previous research. The chapter also includes manipulations of when attention is divided. A thorough review of divided attention and associative deficits is conducted in Chapter 8.

A large meta analysis by Old and Naveh-Benjamin (2008a) examined studies that had separate item and associative memory measures with young and older adults. Ninety studies were analysed and age deficits in associative memory measures were found to be larger than age deficits in item memory measures.

An interesting result from Old and Naveh-Benjamin (2008a) was that although associative memory age deficits were larger than item memory age deficits, there was a significant correlation between the two, $r(88) = .39$, $p < .001$. This indicates that ageing does affect the two different memory abilities similarly to a certain degree. The result provides evidence for a common cause of item and associative memory deficits in older adults. It may be the case that associative deficits are simply cumulative item deficits: To remember an association between items requires memory for the items individually. A numerical example shows how this may work: If young adults have a probability of .9 of remembering an item and older adults have a probability of .7, then item memory age deficits are .2. If remembering an association between two items requires memory for the two items individually then young adults have a maximum probability of .81 (.9 X .9) of remembering the association and older adults have a maximum probability of .49 (.7 X .7) of remembering the association. This gives an age-related associative memory deficit of .32 which is larger than the .2 item memory deficit. Key evidence against this view however, is that older adults have been found to show associative deficits even when item memory is equivalent between young and older adults (e.g., Bastin & Van der Linden, 2005; Kilb & Naveh-Benjamin, 2007; Naveh-Benjamin, Guez, Kilb et al., 2004; Naveh-Benjamin et al., 2009).

In Old and Naveh-Benjamin's (2008a) meta analysis, three factors were considered across the studies analysed, namely, the nature of associative memory and materials memorised by participants (modality, source, context, temporal order, spatial locations and item pairings across verbal and nonverbal material), the instructions given at encoding (intentional versus incidental encoding), and the nature of the memory test (recognition versus recall). With regards to the type of association examined, older adults showed greater associative than item memory deficits relative to young adults for all types of materials except modality. This suggests that memorising the modality of presentation of a stimulus may be less susceptible to cognitive ageing (Chapter 3 investigates associative memory between an object and its modality of presentation in young and older adults). Old and Naveh-Benjamin (2008a) hypothesised that modality information may be encoded more automatically, and is therefore less susceptible to the effects of ageing (cf. Hasher & Zacks, 1979). There was no difference in age-related associative deficits relative to item deficits between verbal and nonverbal materials, even though overall age differences were smaller for verbal material than for nonverbal material. This is possibly driven by the fact that older adults generally have better vocabulary ability than young adults (e.g., Verhaeghen, 2003). Incidental encoding resulted in smaller age deficits than intentional encoding for memory performance in general. Age-related associative deficits relative to item deficits were different across the encoding conditions: Older adults showed larger associative deficits under intentional learning. This supports Naveh-Benjamin's (2000) view that older adults have strategic deficits in that they are less able to apply a memory strategy under intentional learning. Test format affected the difference between age-related associative and item deficits. For tests involving recognition of item memory the associative memory age deficits were larger than item memory age deficits (in support of an ADH) but for tests involving recall of item memory, associative deficits were not much different to item deficits. Old and Naveh-Benjamin (2008a) pointed out that this was partly because when item memory was tested via recall, performance was already low in older adults and this prevented a further drop in performance for associative memory tests. In general, tests involving recall of both item and associative memory showed no significant differences between age-related associative and item memory deficits whereas tests involving recognition of item and associative memory showed larger age-related associative deficits than item deficits.

Dual Process Accounts of Memory

Old and Naveh-Benjamin (2008a) drew links between ADH and dual process accounts of memory. In the previous chapter, dual process accounts of memory were discussed, with age deficits more prominent in measures involving recollection than in measures involving familiarity (e.g., Light et al., 2000; Yonelinas, 2002). Recollection and familiarity are likely to be differentially involved in memory for items and associations. Recollection is seen as a process involved in consciously retrieving specific details and information from memory and is used in source memory (Light et al., 2000), which requires associative memory. Familiarity is seen as simply having a sense that something has been encountered before (Yonelinas, 2002) and is sufficient for completing tests of item memory. This view is expressed in previous research where it has been argued that familiarity-based processes are sufficient to complete an item test whereas associative tests are more reliant on recollection-based processes (Healy, Light, & Chung, 2005; Hockley & Consoli, 1999; Old & Naveh-Benjamin, 2008a). Therefore, a dual process account could explain why age deficits are often smaller for item tests than for associative tests.

The strongest evidence for an ADH (with item and associative measures in the same studies) comes from experiments that use recognition based measures of associative memory but it is still likely that recollection is used to make a response judgement in these tests. In studies such as Naveh-Benjamin's (2000) Experiments 1-3, associative recognition memory is tested for intact and recombined pairs of stimuli. An identification of whether a given pair is intact or recombined cannot be made on the basis of familiarity because both intact and recombined pairs of stimuli contain seen-before items. Therefore, to respond correctly, a participant must recollect specific details about the association (Old & Naveh-Benjamin, 2008a). Furthermore, age deficits in associative recognition memory often arise from older adults' false alarms to lures rather than their hits to targets (e.g., Castel & Craik, 2003; Healy et al., 2005). This suggests that older adults are able to use familiarity to endorse seen-before associations but that they are unable to use recollection to reject recombined associations. Similar findings occur with dual process paradigms involving recognition of associations between items. When items are repeated at study (i.e., familiarity is increased), false alarm rates to lure associations increase in older adults (e.g., Jacoby, 1999; Light et al., 2004) as they are less able to use recollection to reject the highly familiar lures. In contrast, young adults show reduced or similar levels of false alarms upon repetition of items.

Also age deficits are typically smaller for recognition (familiarity based memory) than for recall (recollection based memory) (e.g., Craik & McDowd, 1987; Light et al., 2000; Naveh-Benjamin, 2000; Schonfield & Robertson, 1966). This may be explained by reduced environmental support for recollection tasks compared to recognition tasks. It is therefore consistent with the notion of larger age differences for tasks requiring increased self-initiated processing in the absence of environmental support (e.g., Craik & Byrd, 1982; see Chapter 1 for review). Yonelinas (2002) reviewed evidence showing that recollection is more reliant on prefrontal brain areas than is familiarity. This can be considered alongside age-related decline of prefrontal functionality (West, 1996), which has been linked to age reductions in self-initiated processing (Logan, Sanders, Snyder, Morris, & Buckner, 2002). In the next chapter, age-related associative deficits are considered in terms of distinctiveness. Memory stimuli were made distinct by presenting them in different formats. This produced a highly salient memory stimulus (i.e., increased environmental support) that was designed to be attended to more intensely. This would encourage binding in memory between the stimulus and its presentation format and helped establish if associative deficits occurred when processing was encouraged by distinctiveness. Therefore, the learning of associations was implicit, but participants were strongly influenced to process them, limiting the necessity for self-initiated processing of associative information.

Neuropsychological Deficits

Prefrontal decline in old age has also been linked to source and episodic memory deficits (West, 1996), and may be responsible for age-related associative deficits (Old &

Naveh-Benjamin, 2008a). Moscovitch (1992) proposed that frontal areas of the brain are responsible for strategic use of memory and that they support medial temporal/hippocampal areas which are responsible for explicit, episodic and associative memory. More recently it has also been argued that age-related decline in both of these areas is responsible for episodic/binding memory deficits in older adults (Cabeza, 2006; Mitchell, Johnson, Raye, Mather et al., 2000; Shing et al., 2010). This is supported by a range of evidence such as greater prefrontal activity in young adults relative to older adults during encoding of word pairs (Cabeza et al., 1997), and increased utilisation of prefrontal and hippocampal areas in young adults compared to older adults when binding features to spatial locations compared to encoding locations individually (Mitchell, Johnson, Raye, & D'Esposito, 2000). Increased utilisation of prefrontal areas has also been shown in young adults relative to older adults when encoding pairs of pictures (Iidaka et al., 2001). Cabeza, Anderson, Houle, Mangels, and Nyberg (2000) found greater prefrontal activity in young adults compared to older adults when encoding words within different lists (associating words to a list context). Additionally, increased prefrontal activity has also been shown in young adults when maintaining integrated information (binding of letters and spatial positions) in working memory compared to when they maintained unintegrated information in working memory (Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000).

Outside of neuroimaging, prefrontal lesions have been shown to impact ‘relational’ (associative) memory more than item memory (Cabeza, 2006). Neuropsychological studies have also shown that higher measures of frontal abilities correspond to increased source memory performance in older adults (e.g., Glisky &

Kong, 2008; Glisky, Rubin, & Davidson, 2001). In Chapter 1 it was noted that prefrontal degradation as a result of healthy ageing has been found to occur earlier than degradation of other areas of the brain (e.g., Dennis & Cabeza, 2008; West, 1996). Age-related decline has also been reliably found in the hippocampus (see Raz, 2000, for review). A study by Raz et al. (2005) found substantial shrinkage in hippocampal volume in older adults relative to young adults and longitudinal data over five years showed accelerated shrinkage in this area in older adults relative to young adults.

Two experiments in the thesis were designed to supply insight into the neuropsychological aspects of age-related associative deficits. In Chapter 7, associative memory in children is compared to young adults. Children have been shown to have deficits in forming associations compared to young adults and previous lifespan studies have hypothesised that poor associative memory performance observed in children is due to strategic deficits rather than associative deficits whereas older adults' associative memory is affected by both strategic and associative deficits (Shing et al., 2010; Shing, Werkle-Bergner, Li, & Lindenberger, 2008). Shing et al. (2010) argued that associative memory deficits in children relative to young adults are mainly due to protracted development of the prefrontal cortex and not low hippocampal functionality. Therefore Chapter 7 examined associative memory in children and young adults under high strategic support (low prefrontal requirements) and low strategic support (high prefrontal requirements) conditions. This aimed to establish if children show different patterns of associative deficits to older adults. Chapter 8 also aimed to provide some insight into the role of prefrontal areas in associative memory. It was described earlier how in Chapter 8 young adults' item and associative memory was tested under divided attention; the concurrent task used to divide attention was designed to specifically tax (and therefore disrupt) strategic frontal activity.

Strategy Utilisation Deficits

The impact of prefrontal decline on associative memory in older adults may be a result of deficits in strategy utilisation. It was discussed earlier how Naveh-Benjamin's (2000) Experiment 2 measured strategy use by comparing incidental and intentional associative memory. It was found that incidental learning resulted in a smaller age-related associative deficit than intentional learning, supporting the view that older adults are less able to apply strategies that enhance associative memory performance. This is because the young adults' associative memory performance improved when they were aware of an upcoming memory test but older adults' memory performance did not. A similar result was found in Naveh-Benjamin's (2000) Experiment 3 using different stimuli. Also Naveh-Benjamin (2009) found that intentionally learning the associations between names and faces resulted in larger age-related associative deficits than when participants were asked to simply decide if a name and face 'belonged together'. It was also discussed earlier that in a review of the literature, Old and Naveh-Benjamin (2008a) found that age-related associative deficits were smaller under incidental learning conditions (when strategy has a minimal impact on performance) compared to intentional learning conditions.

Assessments of strategy utilisation in associative memory tests have shown that older adults are less likely than young adults to generate a verbal strategy to link pairs of unrelated pictures (A. D. Smith, Park, Earles, Shaw, & Whiting, 1998) or pairs of unrelated words (Dunlosky & Hertzog, 2001). Strategy utilisation has also been experimentally manipulated in associative memory tasks. Glisky et al. (2001) showed a reduction in source memory deficits in older adults (remembering the room in which a chair was presented) when participants were instructed to think about the relation between the chair and the room. Naveh-Benjamin et al. (2007) showed that when participants were encouraged to use a strategy during learning of word pairs (creating a sentence to link two unrelated words), older adults were able to reduce their associative deficits relative to young adults. Older adults showed a larger difference between item and associative memory than young adults when learning word pairs. This age by test interaction was significantly reduced (i.e., older adults' associative memory improved more than their item memory) when learning occurred under the explicit instruction to use sentences to link the words of each word pair compared to when studying the words normally. The age-related associative deficit was reduced when strategy utilisation was encouraged at encoding and was reduced further when strategy utilisation was encouraged at both encoding and retrieval.

The use of preexisting knowledge may reduce the necessity for controlled and strategic processing in relation to associative memory. When relations between to-be-associated stimuli are easy to comprehend using knowledge acquired before the experimental period, age-related associative deficits have been shown to become smaller. For example, Naveh-Benjamin et al. (2003, Exp. 2) tested young and older adults' recognition of word pairs containing either unrelated or semantically related words. With unrelated word pairs, older adults showed typical associative deficits with larger age deficits for the associations than for the words individually. However, with the semantically related words, the age deficits were similar for associative and individual word memory. This resulted in a triple interaction between age, test type, and pair relatedness. Older adults are therefore able to use preexisting knowledge to improve their associative memory relative to item memory. Attenuation of age-related deficits in associative memory has been found in many studies where preexisting knowledge can be used to support associative memory (e.g., Castel, 2005, 2007; Naveh-Benjamin, 2000; Naveh-Benjamin, Craik, Guez, & Kreuger, 2005; Patterson, Light, Van Ocker, & Olfman, 2009). This literature is reviewed in more detail in Chapter 5 where the amount of preexisting knowledge related to items is manipulated. Furthermore, Chapter 6 manipulates relations between to-be-associated words.

Thesis Overview

The overall aim of the thesis was to evaluate the different theoretical explanations of age-related associative deficits. In general, the main approach used was to find situations where the deficit did not occur. In this sense the thesis also tested the limits of the deficit.

Chapter 3 investigates the binding of items and their method of presentation using encoding conditions favourable to older adults. Assessment of the isolation effect in older adults provided an implicit test of associative memory with a high level of environmental support: Associative memory was encouraged by introducing salient and distinctive ‘isolated’ stimuli amongst less distinctive ‘non-isolated’ stimuli. Additionally, Chapter 4 presents an applied psychology study where the role of age-related associative deficits is considered in relation to eye witness identification of distinctive faces. Again the processing of associations was encouraged via distinctiveness.

Chapters 5 and 6 aimed to elucidate what differentiates an item memory from an associative memory by manipulating preexisting knowledge related to study material. It is hypothesised that the distinction between item and associative memory is that an item memory is an existing structure and associative memories are new links between existing structures. In this sense, an item memory is a re-activation of a unit in memory whereas an associative memory is a novel connection between units in memory. Therefore a memory that is supported by preexisting knowledge requires less associative memory. This definition of item and associative memory is used to test the hypothesis that age-related associative deficits are driven by age deficits in forming new/novel memories. The effect of preexisting knowledge on associative memory is also investigated in children (Chapter 7) who also show associative deficits.

Chapter 8 presents a study that investigates whether associative deficits are caused by global deficits in memory as a result of reduced cognitive resources. In the study, young adults perform item and associative memory tests under full and divided attention. The divided attention condition was specifically designed to tax cognitive resources related to associative memory as the concurrent task was a semantic judgement task which has been shown to activate prefrontal areas. Chapter 8 therefore evaluated whether an associative deficit occurred in young adults who had extra demands on (and therefore reduced cognitive resources in) prefrontal areas. Furthermore, the study aimed to assess which phase of a memory test (encoding or retrieval) was more susceptible to disruption of associative memory by a concurrent task. Item and associative memory were compared under full attention and divided attention at encoding only, retrieval only and encoding and retrieval combined. This would explore the dual process account of associative deficits which is solely based on the retrieval period.

Chapter 9 concludes with a summary of the findings from the thesis and a commentary on the key ideas surrounding age-related associative deficits in light of the current results. This final chapter also considers future directions and discusses potential avenues for ongoing research.

Chapter 3: Age-Related Associative Deficits and the Isolation Effect

The isolation effect occurs when distinctive items of a memory set show a higher probability of being recalled at a later time. For example, a memory set may include a list of words, with one word isolated by being displayed in a different colour. As a control condition in such an experiment, the word that was previously isolated would then be presented without the colour difference in a different trial or to a different participant. In a subsequent memory test, the word in the isolated colour will then show an increased chance of being recalled compared to the same word in the control memory set. This effect is most commonly attributed to a paper by von Restorff (1933) who used a memory set of either nine syllables and one number or the inverse (nine numbers and one syllable).[3] The number or syllable that was distinct/isolated had a higher probability of being successfully recalled; many other experiments have shown similar results with a variety of characteristics determining isolation (see Hunt, 1995; Wallace, 1965, for reviews).

The associative deficit hypothesis developed by Naveh-Benjamin (2000) suggests that older participants have difficulty associating information to and between items. Associative memories are particularly relevant to the isolation effect as the salience of an isolate will be determined by how strongly the isolate is associated to its isolating feature. Older participants should show a reduced isolation effect compared to young participants if they are suffering from associative deficits, as they would be less able than young participants to associate the isolate to the feature that isolates it. If the isolate is not strongly associated to its isolating feature, this would reduce the isolative

[3] Experimental details acquired from a review by Hunt (1995) as the original paper was never printed in English.

properties that produce the superior memory for isolates. A relevant associative deficit measure was included in Naveh-Benjamin's (2000) Experiment 3: After presentation of a memory set, older participants were less able than young participants to recognise associations between words and the fonts they were previously presented in; despite this, they showed equal memory compared to young participants for the words and the fonts individually. This showed that the different features of the stimuli were less likely to be associated in memory by the older participants. The salience of an isolated stimulus is high because it stands out among surrounding stimuli; it may be the case that this salience reduces the necessity for self-initiated processing of associative information and this may reduce age-related differences in associative memory formation (e.g., Craik, 1982, 1986). Therefore, the aim of this chapter is to establish if age-related associative deficits produce age differences in the isolation effect.

Understanding the Isolation Effect

The most obvious place to start when trying to understand the recall benefits of isolated stimuli is to look at the properties of the isolation itself. The isolation effect is present when different characteristics of stimuli determine isolation, for example, colour, size, word type and so on (Wallace, 1965). Therefore it cannot be said that isolation effects are because certain types of stimuli are more beneficial to memory formation. This leaves the distinction of the isolated stimulus from the other list items as the most likely cause of the isolation effect. Green (1956) attributed the isolation effect to surprise, with the isolated stimulus increasing attention and therefore producing superior encoding. This is unlikely when the methodology of von Restorff's (1933) experiment, repeated by Hunt (1995), is considered: The isolated stimuli were intentionally placed at the beginning of the sequentially presented memory set lists so that their distinctiveness would not be apparent at the time of encoding. Therefore, there is no reason to conclude that the distinctiveness of the stimuli evoked a surprise response in participants. In addition to this, Dunlosky, Hunt and Clark (2000) conducted an experiment where participants were asked how likely they were to remember each item of a sequentially presented memory set. When an isolated stimulus was at the beginning of the list, participants rated it no more likely to be remembered than control stimuli. When an isolated stimulus was at the end of the list, it was rated as more likely be recalled than control stimuli. Despite this rating, the magnitude of the isolation effect was not significantly different for both positions of the isolated stimuli, further indicating that surprise cannot account for isolation effects.

To examine the distinction of isolated stimuli further, it is important to avoid circular logic; the definition of the isolated stimulus in an otherwise homogeneous memory set is that it is different or distinct from the non-isolated stimuli. Therefore we gain no insight if we attribute isolation effects to the distinctiveness that defines them. First it must be elucidated what distinctiveness means and then the definition can be applied to the isolation effect. Schmidt (1991) describes distinctiveness as stored representations that lack features of other representations. Schmidt also places emphasis on the importance of a contextual background: For example, in the list ‘mouse, goldfish, piranha, whale’ the mouse is distinctive as it is a land animal but in the list ‘mouse, goldfish, spider, whale’ the whale is distinctive as it is much larger than the other three. The surrounding context is important as it can change the criterion for isolation, as can be seen in the example, which uses two similar lists to produce two different types of isolation. For the purposes of defining the isolation effects, it is sufficient to describe distinctiveness as an incongruity with the surrounding context. For a detailed review of how memory relates to distinctiveness, see Schmidt (1991).

It is hypothesised that distinctiveness results in increased elaboration and rehearsal of the distinct item, therefore increasing the quality of its encoding (Schmidt, 1991). Arguments against this view are similar to arguments against the surprise effect above, namely, that isolation effects are shown to be present with isolated items at the beginning of a memory set before distinction is apparent. The difference here though is that elaboration and rehearsal begin automatically (Schmidt, 1991), producing a similar memory benefit to the primacy effect. To address distinctiveness in the isolation effect, Hunt and Lamb (2001) conducted an experiment whereby participants were required make difference judgements during presentation of a memory set. Participants were required to ‘say out loud something that was different between the item they were viewing and the immediately preceding item’ (p. 1361). When these instructions were followed, isolated items showed no recall benefits above the homogeneous control stimuli. In a further control condition, no distinctiveness judgements were made and the isolation effect returned; in addition, the non-isolated items showed significantly poorer recall. This shows that forcing a judgement of difference can improve memory to the same extent as the isolation effect and therefore leads to the conclusion that the distinctiveness in the isolation effect is similar to marking the differences between stimuli.

Whilst it is most compelling to attribute the isolation effect solely to the level of distinction, there is a body of evidence to show that distinctiveness may not be the only factor that produces recall benefits of isolated stimuli. When looking at isolation effects, it is possible to use a control condition that keeps the target item distinct but non-isolated. If a control list with completely heterogeneous items is used then every item on that list is distinctive. For example, a heterogeneous list may include a noun, a picture, a letter, a symbol and a digit; no isolation occurs as without any homogenous items there is no criterion for item isolation. Isolation effects can therefore be measured by comparing the isolated stimulus (a distinct stimulus on an otherwise homogenous list) to an identical counterpart in a heterogeneous control list. In this situation, the isolation effect is still present even though distinctiveness is equal in experimental and control conditions, as was found by von Restorff (1933) and repeated by Hunt (1995). If direct distinctiveness is not a factor that improves memory of an item, then the isolation effect could be due to interference of the homogenous non-isolated items. In support of this view, Gibson (1940, p. 203) argued that isolation effects are due to '*aggregation* of the traces of the homogeneous items, thereby causing any single item to lose its identity', with encoding of the isolated item not suffering from such aggregation. Watkins and Watkins (1975) showed that similar characteristics of items in a memory set can be proactively inhibited, which reduces the chance of them being later recalled. When combining lists in a memory test, they found that increasing the number of lists within a single category reduced the overall level of recall compared to sets of lists covering multiple categories. It is therefore reasonable to conclude that the isolation effect is partly due to interference during encoding and/or recall among the non-isolated homogeneous items in a memory set.

Isolation Effects in Older Adults

Currently, only five studies exist where isolation effects have been compared between young and older participants (Bireta, Surprenant, & Neath, 2008; Cimbalo & Brink, 1982; Geraci, McDaniel, Manzano, & Roediger, 2009; R. E. Smith, 2011; Vitali et al., 2006). In these studies, isolation effects for older participants compared to young participants were the same (Geraci et al., 2009; Vitali et al., 2006), reduced (Bireta et al., 2008), or completely absent for older participants yet present in young participants (Cimbalo & Brink, 1982); R. E. Smith (2011) found a mixture of age differences within her study. This indicates that isolation effects across age are very sensitive to experimental design, which differed between these experiments.

In the study by Cimbalo and Brink (1982), isolation effects were completely absent in older participants yet present in young participants completing the same memory task. The memory task consisted of a set of nine consonants presented simultaneously, followed by a recall period in which participants wrote down the letters they could recall, writing in the same position as they appeared and guessing if necessary. In the isolation condition, the fifth letter was presented in a larger font (almost double the size of control letters). Although no isolation effect was found in older participants, overall recall performance was significantly better for the isolated condition. After testing, upon questioning of memory strategies, only one of 22 older participants reported awareness of the isolate compared to nine of 35 young participants. The experiment was also conducted with two durations of presentation of the memory set (9 s and 27 s); for the shorter duration, isolation effects were reduced in young participants. To explain a lack of isolation effect in older participants, Cimbalo and

Brink (1982) argued that they are less likely to adopt an encoding structure that does not depend on preexisting strategies. Therefore, they are less likely than young participants to use the isolate as a structural point to aid memory chunking. This is evidenced by the lack of awareness of isolated stimuli for older participants.

Bireta et al. (2008) observed an isolation effect in older adults but found that it was not as strong as the effect in young adults. In their study, a memory set of 12 nouns was presented sequentially. In the control condition all 12 nouns were printed in black and in the isolation condition the nouns were all black apart from the seventh which was printed in red. Memory for the nouns was tested with a free recall test. The majority of both older and young participants were explicitly aware of the isolate upon later questioning. Recall was significantly better for isolates than for the corresponding seventh position controls for both young and older participants; the magnitude of this difference was significantly smaller for older participants. The discovery of an isolation effect in older participants, unlike Cimbalo and Brink (1982), was addressed but no conclusions were reached beyond acknowledging experimental design differences. Bireta et al. (2008) proposed that the reduced isolation effect in older participants is due to a deficit in associating information to the surrounding context. They linked their results to the age-related associative deficits observed by Naveh-Benjamin (2000). This supports the view that older participants are less able than young participants to associate isolates to their isolating features. Therefore, this is perhaps why the older participants showed a smaller isolation effect than the young participants.

The two ageing studies described above are similar in that they both found reduced isolation effects in older participants. In contrast to these, Vitali et al. (2006)

found no difference in isolation effects across age. Vitali et al. (2006) used memory sets of 10 words presented sequentially. To generate an isolation effect, one of the 10 words was displayed double the size of control stimuli; throughout trials, the isolated word was evenly distributed between the fourth and seventh positions inclusively. In the control condition, all words were the same size. At the end of the memory set display sequence there was a 7 s waiting period, then participants wrote down what they remembered (free recall). At the end of the memory test, participants were questioned about awareness of the size difference; all of the 20 young participants and 16 out of 20 older participants had noticed that there were isolated stimuli. The isolation effect was present in both young and older participants, with no significant difference between the two groups.[4] The experiment was also conducted with patients suffering from Alzheimer's disease; their results showed that they had no awareness of the isolate and no isolation effect appeared in their recall data.

Vitali et al. (2006) considered the differences in their results compared to Cimbalo and Brink (1982). They hypothesised that their own results were perhaps due to a reduced difficulty of the memory task, but reached no conclusions about other experimental differences. To explain the lack of isolation effect differences across age, Vitali et al. (2006) considered the possibility of ceiling effects in the young participants. They stated that for young participants the task may have been relatively easy and the improved recall for isolated stimuli could have hit a ceiling. Therefore, the isolation effect may have been greater for young participants if not for the ceiling effect. As older

[4] It is worth noting that Vitali et al. (2006) defined the isolation effect in a slightly different way to the majority of isolation experiments. Instead of representing the isolation effect as the difference in recall between isolates and control stimuli, they represented it as a ratio by dividing the probability of recalling an isolate by the probability of recalling a control. Therefore, a ratio significantly larger than one indicated an isolation effect.

participants had poorer recall overall, they would not have encountered such a ceiling effect for isolated stimuli. The ceiling effect could have reduced the isolation effect in young but not older participants, thus eliminating the difference in isolation effects. This ceiling effect could also explain the differences in age-related isolation effects between Vitali et al. (2006) and Bireta et al. (2008).

Geraci et al. (2009) also found no difference in the isolation effect across age. Young and older participants were presented with memory sets of eight items with semantic category as an isolating factor. Isolates were placed in the fourth, fifth, or sixth position and were of a different semantic category to the rest of the memory set - for example, the word *table* was placed in a list of types of fish. Words were presented sequentially for three seconds each and every participant completed 12 trials. Contrary to the previous studies, Geraci et al. (2009) measured long-term memory for the items: After all 12 lists were presented, participants completed a distracter task for five minutes before undergoing cued recall, which prompted recall with the various categories of controls and isolates. Both young and older participants showed an isolation effect but there was no age difference in isolation effect strength. When questioned about what they noticed about the lists, half of the young and half of the older participants reported awareness of the isolates. For both young and older participants, awareness of the isolate resulted in an isolation effect but for participants unaware of the isolate no isolation effect occurred. Geraci et al. (2009) hypothesised that awareness of isolates could explain the difference between their study and Cimbalo and Brink (1982).

R. E. Smith (2011) found both a presence and an absence of isolation effects in older adults, which depended on experimental conditions. In Experiment 1 a word was isolated based on semantic category. Young and older participants were presented with a single list of eight words and were tested via free recall after a five-minute delay. Isolation lists contained seven words from a semantic category (fish types) and the isolate word ‘table’. Control lists were heterogeneous containing words from eight different categories but with the word table in the same position as the comparable isolation list. Isolation and control lists were a between subjects factor. A crucial manipulation was the position of the isolate. Half of the isolation lists included the isolate as the second word of the list and half as the fifth word of the list. Young adults showed an isolation effect for both isolate positions but older adults only showed an isolation effect for late isolates. R. E. Smith argued that for early isolates, the isolating criteria was not immediately apparent at encoding and therefore provided no environmental support (the distinctiveness of the early isolate is only apparent after it is no longer on screen) to aid older adults’ processing of distinctiveness. In Experiment 2, the isolating factor was designed to be immediately distinctive – isolates were numbers and non-isolates were words. This meant that even though the isolates were early in the list, they were clearly distinctive from non-isolates. This manipulation produced an early isolate effect in both young and older adults and the magnitude of the isolation effect was not significantly different across age group. This experiment indicated that awareness of the distinctiveness of the isolate at the time of its presentation was a factor that altered age differences in the isolation effect.

Although each of the existing studies presents different results, there is one common aspect that seems to affect isolation in older participants. In Bireta et al. (2008), Vitali et al. (2006) and Geraci et al. (2009), older participants showed an isolation effect and in all of these studies the older and young participants were almost equally aware of the presence of the isolate. In Cimbalo and Brink (1982), older participants were less aware of the isolate than were young participants and showed no isolation effect. Also the Vitali et al. (2006) experiment was conducted with Alzheimer's disease patients, who showed no isolation effect and demonstrated no awareness of the presence of isolates. Furthermore, R. E. Smith (2011) showed that older adults only showed an isolation effect when the distinctiveness of the isolate was apparent during its presentation. This indicates that awareness of the isolate could be responsible for the age differences in the existing studies. This is a hypothesis that will be tested with data from Bireta et al. (2008) in the later sections of this chapter. Initially, the exploration of the isolation effect across age was considered in relation to the associative deficit hypothesis.

Experiment 1

Design. The experiment was set up to test the suggestion that the associative deficit hypothesis can be used to explain differences between young and older participants' recall of isolated stimuli, as claimed by Bireta et al. (2008). Three separate criteria for isolation were created: colour, position of item on screen and presentation modality. This arrangement provided separate dimensions for isolation to establish if they differentially affected the level of isolation effect between young and old. In particular, modality was chosen as it has been shown in a meta-analysis by Old and Naveh-Benjamin (2008a) to have no associative binding deficit in older participants. Therefore, if an associative deficit is responsible for a smaller isolation effect in older participants, then using modality as an isolating factor should eliminate age differences in the isolation effect.

Method

Participants. Thirty young adults (23 female), aged 18-24 years (*M* = 19.2, *SD* = 1.6), and 30 healthy older adults (21 female), aged 57-88 years (*M* = 73.8, *SD* = 8.0) took part in the experiment. Young participants were first year psychology undergraduates at Warwick University (UK) who participated in exchange for course credit. Older participants were recruited from the University of Warwick Age and Memory Study volunteer panel that was populated by local advertisements; they were offered no financial incentives for participation.

To assess cognitive functioning, participants completed the Digit Symbol Substitution task (Wechsler, 1981) as a measure of perceptual-motor or processing speed. They also completed the multiple choice part of the Mill Hill vocabulary test (Raven, Raven, & Court, 1988) as a measure of crystallised intelligence (see Appendix 1 for more details of these measures). This demonstrated that the participants used in this study had similar cognitive ability and age differences to typical participants in the literature (e.g., Horn & Cattell, 1967; Salthouse, 1991). Young participants were significantly faster at the digit symbol substitution task than older participants, $t(58) = 9.98$, $p < .001$ (young *M* = 74.9, *SD* = 10.7; older *M* = 43.4, *SD* = 13.6), demonstrating that young participants had faster cognitive processing than older participants. For the vocabulary test, young participants scored significantly lower than older participants, $t(58) = 5.17$, $p < .001$ (young *M* = 16.3, *SD* = 3.1; older *M* = 21.3, *SD* = 4.4), demonstrating that young participants had poorer vocabulary than older participants.

Stimuli. Two hundred and forty nouns were taken from Bradley and Lang (1999). The nouns were selected for medium to high valence (*M* = 5.76, range of 2.80-8.56 out of possible 1-9) and for between 5-7 letters in length (*M* = 5.89, *SD* = 0.77). The frequency of the words averaged 8.76, with a range of 6.05-11.65, using log HAL frequency (Lund & Burgess, 1996).

Each trial in the experiment consisted of the sequential presentation of 12 nouns. A control condition was used which presented 12 separate nouns sequentially, each for a duration of 1500 ms. The words were presented in the centre of a computer screen at a viewing distance of approximately 50 cm in black text on a white background with a height corresponding to approximately 1.5º viewing angle.

Three separate conditions were used to create isolation effects; in the isolation conditions every item was presented in the same way as the control condition apart from the isolate, which was always presented as the seventh item in the sequence of 12. The criteria for isolation were as follows: Colour – the seventh item was presented in red text; Position – the seventh item was presented vertically off centre, 55 mm (~ 6.2º) above the non-isolated words; Modality – the seventh item was presented auditorily through headphones in a male voice. In this last case, at the same time as the spoken word, a set of five hash symbols (#) were presented for 1500 ms. This created a visual cue which helped to indicate that the word was part of the experiment. As all non-isolated items were presented in identical modes, the control stimuli were appropriate to all isolate conditions.

Procedure. The experiment involved four conditions – three isolation conditions (colour, position and modality) and one control. Twenty trials were presented to each participant with five trials from each of the four conditions; the order of the trials was completely randomised. Words were presented sequentially for 1500 ms each with no interval between them. All 240 words were presented to every participant once each (20 sets of 12 stimuli). Words were completely randomised so that any word could occur in any trial, position, condition or state of isolation. After each trial of 12 words, participants were required to say all the words they could remember in any order. The remembered words were written down in order of recall on a sheet by the experimenter. Participants were given as much time as they needed to recall the words and were able to rest between trials before pressing a button to continue.

All participants were tested on the same laptop computer and completed two practice trials before the experiment. Three of the young participants were tested in their own homes and 27 at Warwick University in a quiet room; 28 of the older participants were tested in their own homes and two were tested at the university.[5] In the first practice trial, all words were presented in the same way as controls (all practice words were different to experimental words). In the second practice trial, each of the isolation presentation types was presented at least twice in one sequence of 12 words; participants were made explicitly aware of the different ways in which words would be presented. Upon starting the experimental memory test participants were instructed that *some* words would be presented in these different ways and that ‘you should attempt to

[5] The testing environment differed for young and older participants. This was deemed as acceptable as young (psychology students) and older participants differ in familiarity with laboratory conditions. Therefore, using controlled laboratory conditions for both groups would not have affected young and older participants in the same way. No literature was found examining the effects of laboratory conditions on older participants; mixed results were found for age differences between controlled and naturalistic tasks. Some experiments show age-related decline for naturalistic tasks (see Light, 1991, for examples). Sometimes naturalistic tasks eliminate age differences (e.g., Garden, Phillips, & MacPherson, 2001) and sometimes naturalistic tasks show age-related improvement (e.g., Rendell & Thompson, 1999).

remember any words you see or hear regardless of how they are presented’. It was not mentioned that only one word in each list would be presented differently.

Results

Word recall. Figures 1-3 show the mean proportion of recalled words for serial positions 1-12 of the word lists, comparing control lists to position, modality and colour isolation lists, respectively. The isolation effect is most obvious in the modality lists (Figure 2) and there is a clear perturbation caused by isolates in the position and colour isolation lists. It is clear from the figures that there are primacy and recency effects in the recall data. It is also apparent between the figures that the isolation effect is different in magnitude for the different isolation criteria. Overall the isolation effect is similar for young and older adults, indicating that older adults were successfully able to associate the isolates to their isolating factors.

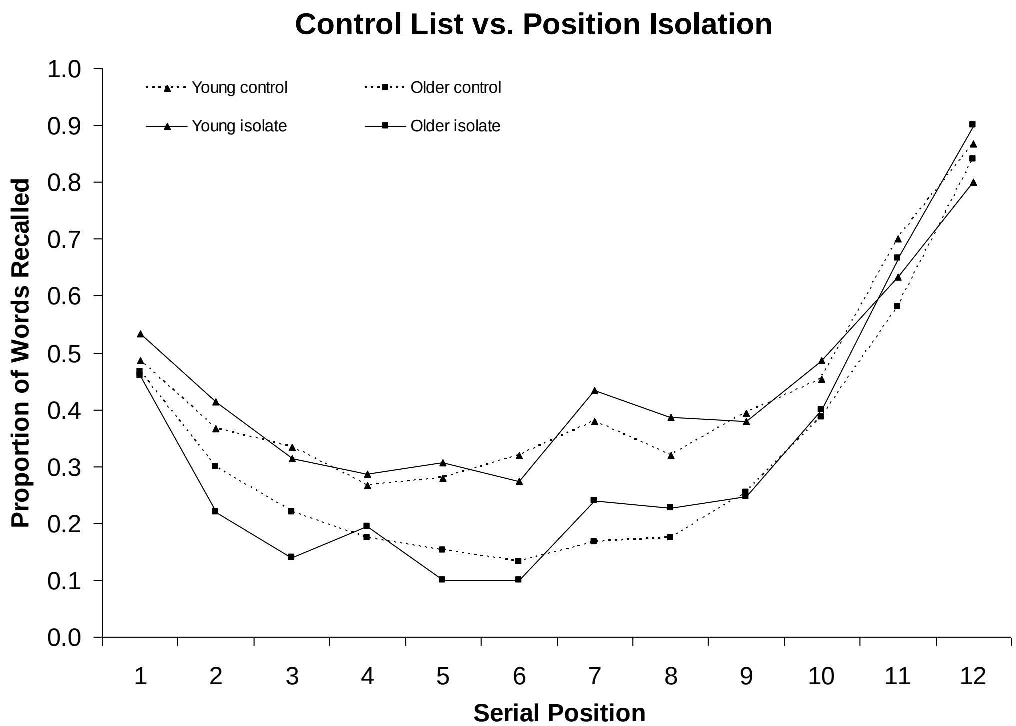

Figure 1. Mean proportion of words recalled as a function of serial position for control and position isolation lists.

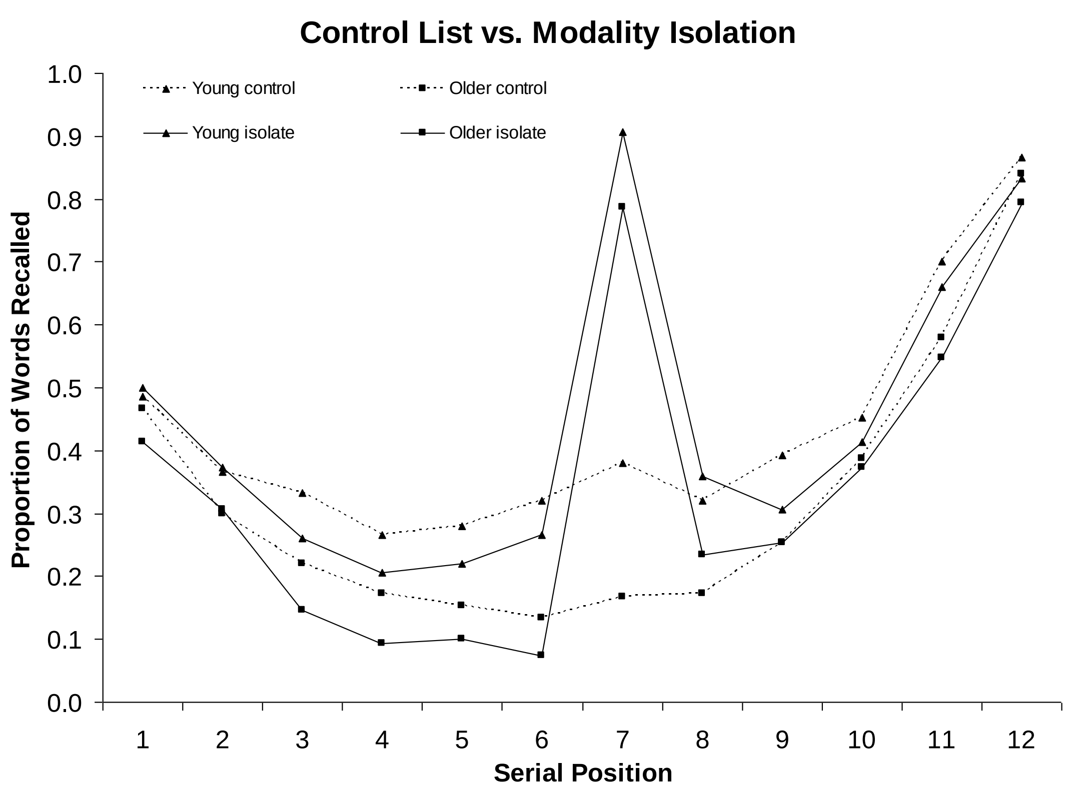

Figure 2. Mean proportion of words recalled as a function of serial position for control and modality isolation lists.

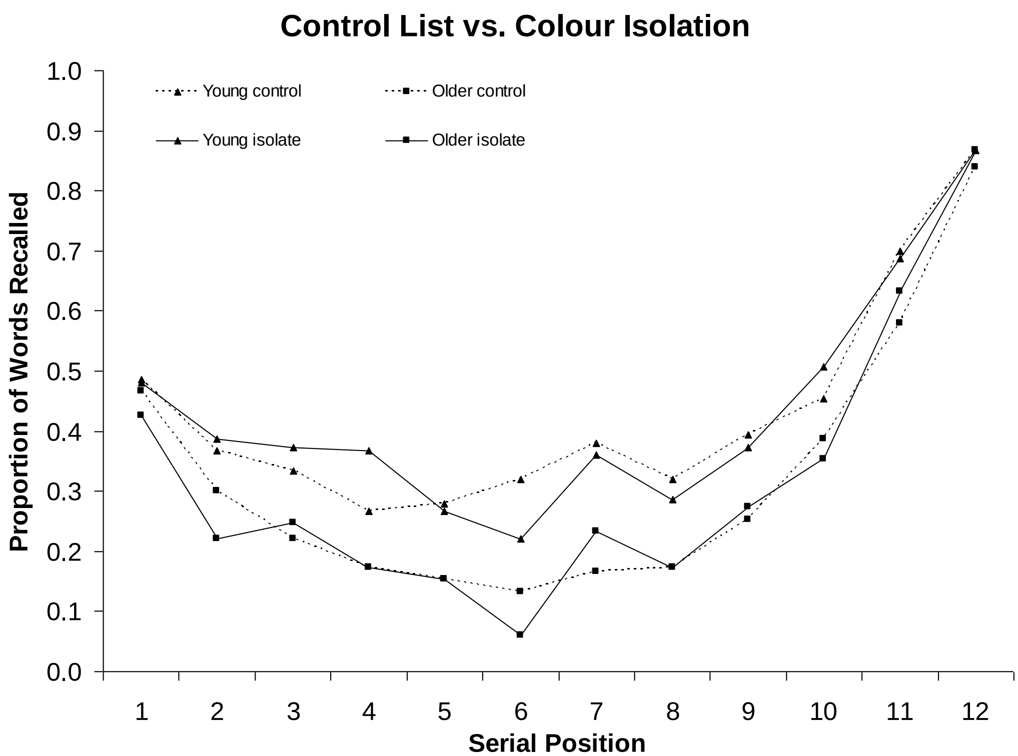

Figure 3. Mean proportion of words recalled as a function of serial position for control and colour isolation lists.

Initially, a 2 (Age: young, older) x 4 (List type: control, position isolation, modality isolation, colour isolation) x 12 (Serial position: 1-12) repeated measures ANOVA was conducted on the probability of correct recall. Mauchly's test indicated that there were violations of sphericity: For serial position, $\chi^2(65) = 320.40$, $p < .001$, therefore serial position degrees of freedom (and consequently Serial position x Age) were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .32$). For Serial position x List type, $\chi^2(560) = 633.51$, $p < .05$, therefore Serial position x List type degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .60$). Unless stated otherwise, all further results showed no violation of sphericity.

There was a main effect of age, $F(1, 58) = 34.10$, $MSE = 0.25$, $p < .001$, with young participants recalling more words on average than older participants (*M* (*SD*) =

5.22 (3.56), and 3.93(3.77), respectively). There was a main effect of serial position, $F(3.57, 206.85) = 110.93$, $MSE = 0.26$, $p < .001$, which showed the presence of primacy and recency effects. There was no interaction between serial position and age, $F(3.57, 206.85) = 1.89$, $MSE = 0.26$, *ns*, indicating that young and older participants displayed similar primacy and recency effects. There was an interaction between list type and serial position, $F(19.83, 1150.17) = 12.77$, $MSE = 0.06$, $p < .001$, suggesting the presence of an isolation effect. There was no main effect of list type, $F(3, 174) = 1.72$, $MSE = 0.03$, *ns*, which demonstrated that the isolates did not change overall list recall. All other interactions were non-significant, $F < 1$.

Isolation effect. To assess the isolation effect, a 2 (Age: young, older) x 4 (List type: control, position isolation, modality isolation, colour isolation) repeated measures ANOVA was conducted on the seventh item of each list type. There was a main effect of age, $F(1, 58) = 22.72$, $MSE = 0.07$, $p < .001$, and a main effect of list type, $F(3, 174) = 104.97$, $MSE = 0.04$, $p < .001$, but no interaction, $F < 1$.

To determine which presentation types caused isolation effects, *t*-tests were performed to compare the difference between recall of the control position seven words and each of the isolate types. Only modality isolation showed a significant isolation effect (young: $t(29) = 10.88$, $p < .001$; older: $t(29) = 12.60$, $p < .001$) but the modality isolation effect was not significantly different between young and older groups, $t(58) = 1.35$, *ns*. The modality isolation effect was so extreme that many participants recalled all of the modality isolated words. All five of the modality isolated words were recalled by 19 (63%) of the young participants and 11 (36%) of the older participants. This ceiling effect may have eliminated any isolation effect differences across age as both age groups performed highly.

It can be seen in Figures 1-3 that words adjacent to isolates showed a slightly different pattern of recall to isolates and equivalent controls. As a different measure of the isolation effect, seventh position word recall performance was compared to the average of the sixth and eighth position word recall performance in paired *t*-tests. Comparisons were conducted for control, position isolation, modality isolation and colour isolation lists separately and for young and older participants. In line with expectations, control lists showed no difference between seventh position word recall and sixth and eighth averaged (young, $t(29) = 1.52$, *ns*, older, $t < 1$). This measure of isolation yielded a marginal isolation effect for position isolates (young, $t(29) = 1.87$, $p = .07$, older, $t(29) = 1.94$, $p = .06$), there was clear evidence of modality isolation (young, $t(29) = 15.99$, $p < .001$, older, $t(29) = 13.75$, $p < .001$), and the new measure demonstrated a significant colour isolation effect (young, $t(29) = 2.32$, $p < .05$, older, $t(29) = 2.48$, $p < .05$). There were no significant age differences with the new measure of isolation (all *t*s < 1). Overall, isolation effects were generally present and importantly the isolation effects were of similar magnitude in young and older adults.

Order of recall. To further investigate the causes of the isolation effect, analysis was conducted on the output position in which the isolate was recalled. Figure 4 shows the mean positions in which the seventh item or isolate was recalled for each list type and age group (mean positions when the seventh item was recalled were averaged for each participant then each participant’s mean was represented with equal weight). From the figure, it can be seen that modality isolates were recalled earlier on average than controls for both young and older participants.

A 2 (Age: young, older) x 4 (List type: control, position isolation, modality isolation, colour isolation) repeated measures ANOVA was conducted on the output positions of the seventh items/isolates. There was a marginal main effect of age, $F(1, 29) = 2.97$, $MSE = 3.25$, $p < .10$. There was a main effect of list type, $F(3, 87) = 7.65$, $MSE = 1.10$, $p < .001$, due to an earlier modality isolate's recall output position, but no significant interaction, $F(3, 87) = 1.05$, *ns*.

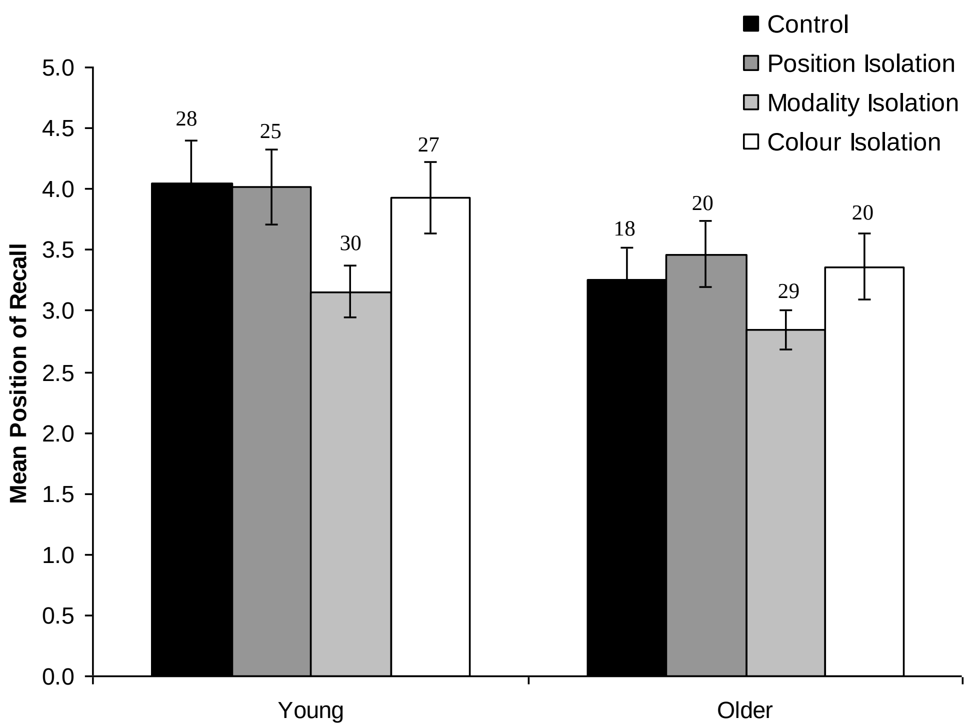

Figure 4. Mean position of recall for the seventh item/isolate of each list type. Error bars are ±1 *SE*. Numbers above the bars indicate the number of participants represented by each bar (some participants did not recall any seventh position words for certain list types).

Intrusions. Intrusions during recall were grouped into four categories based on where the intruded word came from: intrusions from previous list, intrusions from two lists ago, intrusions from three or more lists ago and intrusions from words not in the experiment.[6] Mean intrusions per list were averaged together regardless of list types and are plotted in Figure 5. A series of *t*-tests showed that young participants had significantly fewer intrusions than older participants for all categories (see Table 2). For young and older participants, the differences between the types of intrusions were similar, with the majority of intrusions from within the experiment coming from the immediately-preceding list.

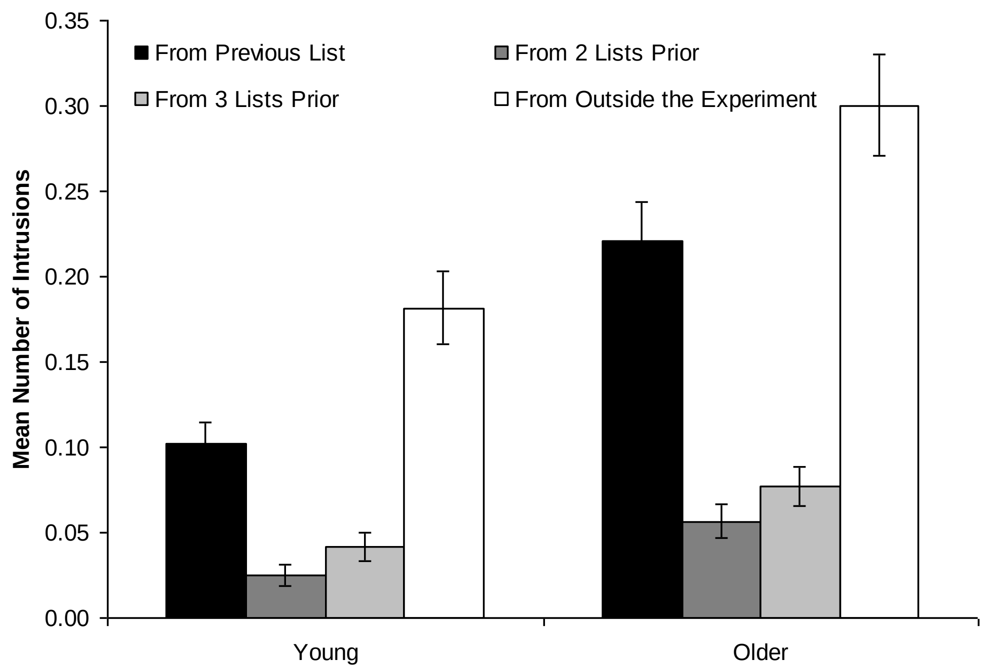

Figure 5. The mean number of intrusions from the previous list, from two lists prior, from three or more lists prior and from outside the experiment. Data are shown for both young and older age groups and are averaged across list types. Error bars are ±1 *SE*.

[6] Occasionally a word from the experimental word list was stated by the participant before they had seen it at all. In this situation, the intrusion was categorised as an intrusion from outside of the experiment.

Table 2

Age Differences in the Number of Intrusions Reported by Participants During the Recall Period

	Age differences
Intrusion type	*t*(58)
From previous list	3.73***
From two lists prior	2.39*
From three or more lists prior	2.14*
From outside of the experiment	2.14*
Total intrusions	3.45**

Note. Positive *t*-values indicate higher intrusions for older participants compared to young participants.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Awareness. To establish how awareness (and also possibly fatigue) impacts upon the isolation effect for different age groups, the experimental data were split into two halves. The probability of recalling the seventh item of the control and isolate lists was calculated separately for the first and second halves of the experiment. If awareness and knowledge of the structure of the lists increase the isolation effect, then the second half of the experiment may show a greater isolation effect as participants become familiar with the word lists.

Table 3 shows the mean probability of recalling seventh position words for each half of the experiment, from each list and for young and older participants. Isolation effects were calculated by subtracting the probability of recalling the seventh control word from each half from the corresponding probability of recalling each isolate type from the same half. A 2 (Age: young, older) x 2 (List half: first half, second half) x 3

(Isolation type: position isolation, modality isolation, colour isolation) ANOVA was conducted on the recall data. There was no main effect of age, $F < 1$. There was a main effect of isolate type, $F(2, 102) = 85.87$, $MSE = 0.11$, $p < .001$, which was not surprising as modality isolates were much more likely to be recalled than any other. Crucially, with regard to familiarity and awareness, there was a marginal effect of experimental half, $F(1, 51) = 3.26$, $MSE = 0.30$, $p = .08$, with greater isolation effects for the second half. There were no significant interactions (all *F*s < 1.82).

To look at how the recall of seventh items for different list types varied between the halves, *t*-tests were performed between the halves separately for young and older participants and for each list type (see Table 3). None of the *t*-tests showed significant differences between the halves except for the control lists of young participants. For young participants, significantly more words from the seventh position in the control list were recalled from the first half of the experiment than the second, $t(28) = 2.25$, $p < .05$. This could partly account for the increased isolation effect in the second half of the experiment. The small local peak for the seventh word from the young participants' control lists is visible on Figures 1-3 and it is now known that this comes mainly from earlier trials.

Table 3

Mean Probability of Recall of Seventh Position Control and Isolate Words for the First and Second Halves of the Experiment and for Young and Older Participants

	First half		Second half		
List type	*M*	*SD*	*M*	*SD*	*t*
		Young			
Control	.45	.37	.31	.25	2.25*
Position	.42	.38	.42	.35	0.00
Modality	.89	.25	.92	.19	-0.40
Colour	.30	.32	.42	.37	-1.24
		Older			
Control	.18	.29	.18	.23	0.00
Position	.27	.32	.26	.33	0.15
Modality	.73	.39	.82	.29	-1.02
Colour	.20	.30	.28	.31	-1.08

Note. *t*-values show differences of means between halves.
*$p < .05$.

Discussion

The experiment successfully generated an isolation effect for both young and older participants with modality isolated words, and there was also some evidence of an isolation effect for position and colour isolated words. Contrary to some studies in the literature, there were no significant differences in the isolation effect between young and older participants. Other aspects of the data showed age differences, particularly overall recall, which was lower for older participants as one would expect (e.g., see Zacks et al., 2000). Also, older participants were more likely than young participants to incorrectly recall an intrusion, consistent with the inhibitory deficit theory of ageing (Hasher & Zacks, 1988).

Isolation effect. The modality isolation effect was very clear for both young and older participants. Words that were isolated by presentation modality (played through headphones instead of presented visually) were more likely to be successfully recalled than non-isolated control words in the same position in the list. There were no age differences in the strength of modality isolation although recall of the auditorily presented isolates was so successful that ceiling effects could have minimised age differences. The colour and position isolates were not recalled significantly better than controls. However, when words adjacent to isolates were compared to isolates, there were some isolation effects for colour isolates and marginal isolation effects for position isolates (equivalent for young and older adults).

The reason that modality isolates produced such a strong isolation effect is probably due to the different way in which visual and auditory information is processed. The three-component model of working memory proposed by Baddeley and Hitch (1974) makes a clear distinction between visual and phonological slave systems. In the model, a Visuo-Spatial Sketchpad is responsible for maintaining visual and spatial information in working memory and a Phonological Loop is responsible for maintaining verbal information in working memory. Through interference-based experiments, the Visuo-Spatial Sketchpad and the Phonological Loop have been dissociated.

Phonological information interferes with other phonological information (e.g., the phonological similarity effect; Conrad, 1964a, 1964b) and spatial information interferes with visual information but not verbal information (Baddeley, Grant, Wight, & Thompson, 1973). Also phonological information has been shown not to interfere with visual items in memory; Murray (1968) showed that acoustic confusability (where acoustically similar verbal material in memory tests is confused upon recall) did not affect recall of visually presented acoustically confusable words when the phonological loop was blocked by articulatory suppression.

In the case of the modality isolation condition of the experiment, the auditorily presented isolate would have gained direct access to the Phonological Loop whilst the visually presented non-isolates (and control words) would have been processed initially by the Visuo-Spatial Sketchpad. This means that the interference between the modality isolate and other list items in working memory would have been reduced. As mentioned in the introduction to this chapter, the interference between non-isolates is a contributory factor to the isolation effect. Therefore recall of the modality isolates may have been significantly more successful than equivalent control words because they suffered less interference from surrounding items.

An alternative explanation for the strong modality isolation effect concerns rehearsal of the words in working memory. Verbal material is believed to be rehearsed as phonological information even when initially presented visually; this is evidenced by a range of phonological experiments (see Baddeley, 2003, for review). The auditorily presented isolate may have had increased probability of recall because it was presented directly in a phonological form without having to be transferred from a visual to a phonological form for rehearsal. That is, the isolate may have been more efficiently rehearsed than non-isolates.

An additional factor that could have benefited modality isolates over colour and position isolates is the modality match effect (see Mulligan & Osborn, 2009, for review). This is where superior memory occurs when the study and test modality matches. In the current experiment, recall was verbal and therefore the modality matched with respect to the modality isolate. This could possibly enhance recall of modality isolates even further compared to the other isolation types and control stimuli. Finally, the order of output may also have enhanced recall of modality isolates. Modality isolates were generally output early in free recall; therefore, they would suffer from less output interference.

There were no strong isolation effects for position or colour isolation lists. For the colour lists, this appears to be inconsistent with Bireta et al. (2008), who found a clear colour isolation effect. The main differences between this study and Bireta et al. (2008) are the presence of other isolation types during the experiment and the reduced number of control lists. With Experiment 1, only 25% of lists were control lists whereas for Bireta et al. (2008), 50% of lists were controls. There could have been increased anticipation of an isolate in this study as the majority of lists contained isolates; it can be seen in Figures 1-3 that for the control lists there is a small local peak in position seven for young participants. This peak could be due to anticipation of the isolates although paradoxically the number of seventh word controls recalled by young participants was significantly larger for the first half of the experiment. If anticipation caused the peak it would be expected to occur more in the second half of the experiment once participants became familiar with the list structures. Alternatively, it could be possible that the young participants were quick to presume isolates in the seventh position of every list but revised their expectations after seeing several control lists. These observations based on anticipation are consistent with the isolation study by Detterman (1975, Exp. 1): recall of control words in the same position as isolates was improved for participants who were knowledgeable about the list structure and proportion of isolate/control lists, compared to uninformed participants. Further comparisons with Bireta et al. (2008) can be seen in the later sections of this chapter.

A final point to consider that could explain the small isolation effects for colour and position lists is the amount of isolation. Gumenik and Levitt (1968) manipulated the amount of isolation by changing the size of isolates. They found that as isolates differed more from controls in size the isolation effect increased. The colour and position isolates may simply not have been sufficiently different from controls to elicit an isolation effect. This is not supported by the success of colour isolation in many experiments (e.g., Bireta et al., 2008; Jones & Jones, 1942; M. H. Smith, 1949). The same cannot be said for spatial/position isolation as, remarkably, no literature was found where spatial position of stimuli was an isolating factor (although one study showed memory improvements for the spatial position of objects isolated by colour; Guerard, Hughes, & Tremblay, 2008). The important point to consider here is that because the modality isolation effect was so strong, the position and colour isolates were not as strongly isolated in relative terms in the context of the mixed isolation experiment. Therefore, perhaps participants reacted to the amount of isolation on a single dimension with position, colour and modality isolation represented on a single internal scale. This means that the strong modality isolation effect may have reduced the colour and position isolation effects.

Effect of an isolate on non-isolates. The results showed that there was no effect on overall recall of a list when isolates were present, which seems to be in general agreement with the literature. This pattern of results was found in the ageing study by Bireta et al. (2008) where overall list recall was not significantly different between isolation and control lists. In contrast, the ageing study by Cimbalo and Brink (1982) found improved overall memory for isolation lists compared to controls. This occurred for both young and older groups even though the older participants showed no isolation effect. Cimbalo and Brink (1982) used simultaneous presentation of their memory sets with the isolates in a central position. They argued that the isolates provided a structure to the isolation lists, which improved overall memory performance. Smith and Stearns (1949) concluded that the presence of an isolate can improve overall learning of a list as it aids in structural organisation but that generally overall list learning is usually equivalent between isolation and control lists. In the Smith and Stearns (1949) experiment, the improved recall of an isolate was accompanied by reduced recall of non-isolates compared to control lists. In a review of the isolation effect, Wallace (1965) concluded that most studies show no overall list recall benefit for isolate lists compared to control lists. More recent studies also show the same pattern of results (e.g., Hunt & Lamb, 2001; Kelley & Nairne, 2001).

From Figures 1-3, it can be seen that words preceding the isolate were less likely to be recalled than control words. During the experiment, several participants reported that when the isolate appeared, they forgot the words preceding it. This indicates that the appearance of the isolate may disrupt rehearsal processes in working memory. In Figures 1 and 2, the word in the eighth position immediately after the isolate appears more likely to be recalled than the equivalent control word. Analysis revealed that isolates were recalled significantly more often than adjacent words for colour and modality isolates and marginally more often for position isolates.

In the literature, there are mixed results concerning recall of items adjacent to the isolate. M. H. Smith (1948) found improved recall of words adjacent to a colour isolate, but later found different results using similar procedures of colour isolation: In agreement with the current study, results from M. H. Smith (1949) indicated a reduced recall of words immediately preceding a colour isolate compared to controls. Unfortunately, this aspect of the data was not statistically tested. M. H. Smith and Stearns (1949) looked at words adjacent to a colour isolate and found little difference between the word preceding the isolate and control words. However, they did find an improvement of recalling the word immediately after the isolate compared to controls. Also, in their figures, Jones and Jones (1942) showed slight recall benefits to items adjacent to a colour isolate but these observations were not statistically tested.

Intrusions. The number of intrusions showed clear age differences, with older participants producing more intrusions than young participants. This is consistent with the reduced inhibition hypothesis, where older participants show less ability than young participants to inhibit irrelevant information (Hasher & Zacks, 1988; Zacks & Hasher, 1997). Following this hypothesis, the words from previous lists were therefore less likely to be inhibited successfully by older participants than young participants, which resulted in increased intrusions. Intrusions from outside of the experiment for older participants could also be explained in terms of reduced inhibition. It is possible that extra-experimental words were selected during retrieval attempts and that older participants were less successful than young at inhibiting them before reporting them.

The increased amount of intrusions for older participants is also consistent with the ageing literature surrounding the Deese-Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). The DRM paradigm is designed to encourage intrusions by creating false memories. Lists of items are presented with a common theme (e.g., glass, curtain, view and frame) which are all associated to a lure word that was never presented (e.g., window). After viewing such lists, when memory is tested, participants are highly likely to incorrectly report the presence of the lure in recognition and recall tests (Roediger & McDermott, 1995). Older participants are more likely than young participants to report the lure (see Schacter, Koutstaal, & Norman, 1997, for review) indicating their increased likelihood of generating intrusions.

The increased number of intrusions of words from within the experiment is congruent with the associative deficit hypothesis developed by Naveh-Benjamin (2000). The hypothesis states that age-related memory deficits are due to increased difficulty for older participants when creating memories that associate items to their different contexts. Thus, there are many examples in the literature that show age-related deficits in memory for the source of information. Older participants have been shown to have poorer memory for the font, colour, case, and sex of voice that words were presented in compared to young participants(see Zacks et al., 2000, for review). In the context of the increased intrusions for older participants, it is apparent that the older participants were less able than young participants to remember when an item was presented. They were less able to associate the words presented to the temporal context in which they were presented and were more likely to report an irrelevant word from an earlier period in the experiment.

Output order. Modality isolates were reported significantly earlier than control words during recall for both young and older participants (see Figure 4). The same was not true for position and colour isolates. As modality isolates were the only isolates to generate a clear isolation effect, there appears to be a relationship between the isolation effect and the output position during recall. In contrast to these findings, Bireta et al. (2008) found that isolates were not recalled significantly earlier than controls. In agreement with the early output of modality isolates, Lewandowsky, Nimmo and Brown (2008) found that temporally isolated letters were output earlier during recall. In addition to this, Lewandowsky et al. (2008) also found that forcing participants to output their response sequentially eliminated the isolation effect. The same was found by Parmentier, King and Dennis (2006); no temporal isolation effects occurred for auditory-verbal and spatial items when output order was restricted to sequential output. These results indicate that the recall benefits for isolated stimuli are related to the priority in which they are later output. If isolates are likely to be output earlier then there is less chance they will have left short-term memory. Also, if isolates are recalled earlier, they will suffer less interference from other items that are output during the same recall period.

Summary. Overall, the results did not reproduce age differences in the isolation effect that is present in some studies. It is apparent that because no age-related isolation effect differences were found, age-related associative deficits may not be responsible for age differences in the isolation effect. This conclusion is reasonable as age-related associative deficits are widely and reliably found in the literature (see Old & Naveh-Benjamin, 2008a; Spencer & Raz, 1995, for reviews). This means that the older participants in the study were almost certainly suffering from associative deficits, yet this did not impact upon their ability to show an isolation effect. The same argument can be applied to Geraci et al. (2009) and Vitali et al. (2006) where no age differences were present in the isolation effect. It is therefore apparent that associating in memory the isolate to its isolating feature is not the main factor that alters the strength of the isolation effect.

It is prudent to note that the results from the current study showed no age differences for modality isolates. Modality is a domain where age-related associative deficits are typically not present (Old & Naveh-Benjamin, 2008a). This finding is therefore not in disagreement with the hypothesis that age-related associative deficits cause age-related isolation effect differences (Bireta et al., 2008). However, there were no age differences for colour or position isolation effects either so this reduces the possibility of drawing firm conclusions from this result.

Aside from the isolation effects, age differences were apparent in overall recall and number of intrusions, with poorer recall and increased intrusions in older participants compared to young participants. These findings were consistent with the general ageing literature (Zacks et al., 2000). Also the presence of isolates seemed to be generally detrimental to adjacent words, suggesting that isolates disrupt rehearsal mechanisms in both young and older participants.

With respect to associative deficits, older participants show larger associative deficits than young participants when memory for associations is tested explicitly (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008a). To clarify, when participants are informed that they must try to remember associations, the age-related associative deficits increase compared to when participants are given a surprise associative memory test. The associative memories created during the isolation effect (when associating the isolate to the isolating feature) are more implicit than explicit; therefore this may be why associative deficits do not consistently alter the isolation effect strength. Instructing participants to particularly memorise the isolate would bring the isolation effect into the explicit domain but the nature of such a task would produce ceiling effects as it would be very easy to memorise the single isolate. Additionally, as the isolates were distinct they may have provided older adults with environmental support at encoding, which has been shown to alleviate the associative deficit (e.g., Naveh-Benjamin, Hussain et al., 2003 and see Chapter 2). R. E. Smith (2011) also demonstrated that older adults did not show an isolation effect when distinctiveness of an isolate was not apparent during encoding. In the next part of this chapter, the effect of awareness of the isolate will be considered in order to ascertain whether awareness of the isolate is responsible for age differences in the isolation effect.

The Isolation Effect Across the Experimental Period

A factor that may influence the isolation effect is awareness of the isolate. As mentioned in the introduction to this chapter, when older participants were aware of the isolate they produced an isolation effect (Bireta et al., 2008; Geraci et al., 2009; Vitali et al., 2006) yet when they were largely unaware they did not (Cimbalo & Brink, 1982; R.

E. Smith, 2011). Measures of awareness were all taken after the experimental period; it is therefore feasible that awareness of the isolate did not necessarily occur from the onset of the experiment. This is particularly relevant to Bireta et al. (2008), where age differences in the isolation effect were present. Older participants may have taken longer than young participants to become aware of the isolate. They may have only produced an isolation effect for the proportion of the experiment when they were aware of the isolate and hence showed a smaller average isolation effect than young participants. If this was the case, it would be expected to see a different isolation effect size for older participants between the first and second halves of the experiment. The raw data from Bireta et al. (2008) were obtained and were re-analysed in order to establish the isolation effect differences across the first and second halves of the experiment.[7]

Procedure. A brief outline of the Bireta et al. (2008) method will be provided here for clarity. A comprehensive account can be found in the original paper. The isolation effect was determined for groups of young (*M* = 19.3 years) and older (*M* = 70.1 years) participants. Memory sets of 12 nouns were presented sequentially. Two presentation rates were used between participants: fast (1500 ms per word) and slow (3000 ms per word). In the control condition, all 12 nouns were printed in black and in the isolation condition the nouns were all black apart from the seventh which was printed in red. Memory for the nouns was tested immediately after the memory set presentation with a written free recall test. Each participant completed 20 trials, half of which were control lists and the other half contained an isolate. Trial types were randomised across the experimental period.

[7] The author would like to offer special thanks to Dr Tamra Bireta for sharing the raw data and providing clear notes outlining the various fields in the data files.

Results

Comparison. Initially, a direct comparison was made between Experiment 1 above and the Bireta et al. (2008) data. The colour and control conditions were extracted from Experiment 1 and plotted against the fast colour isolation experiment from Bireta et al. (2008). Both sets of data used the same sequential presentation rate of 1500 ms per word, both used a red colour isolate in the seventh position of a 12 word list and both used free recall to measure memory. Figure 6 shows the comparison of control list data between the two experiments. Figure 7 shows the comparison of the isolation list data between the two experiments. In general, performance was higher for Bireta et al. (2008) than for Experiment 1. Also recency effects were larger in Experiment 1 than in Bireta et al. (2008), presumably because of the different recall methods (written vs. spoken, respectively). The isolation effect was also larger in Bireta et al. (2008) than Experiment 1. A statistical comparison of the experiments can be found in Appendix 2.

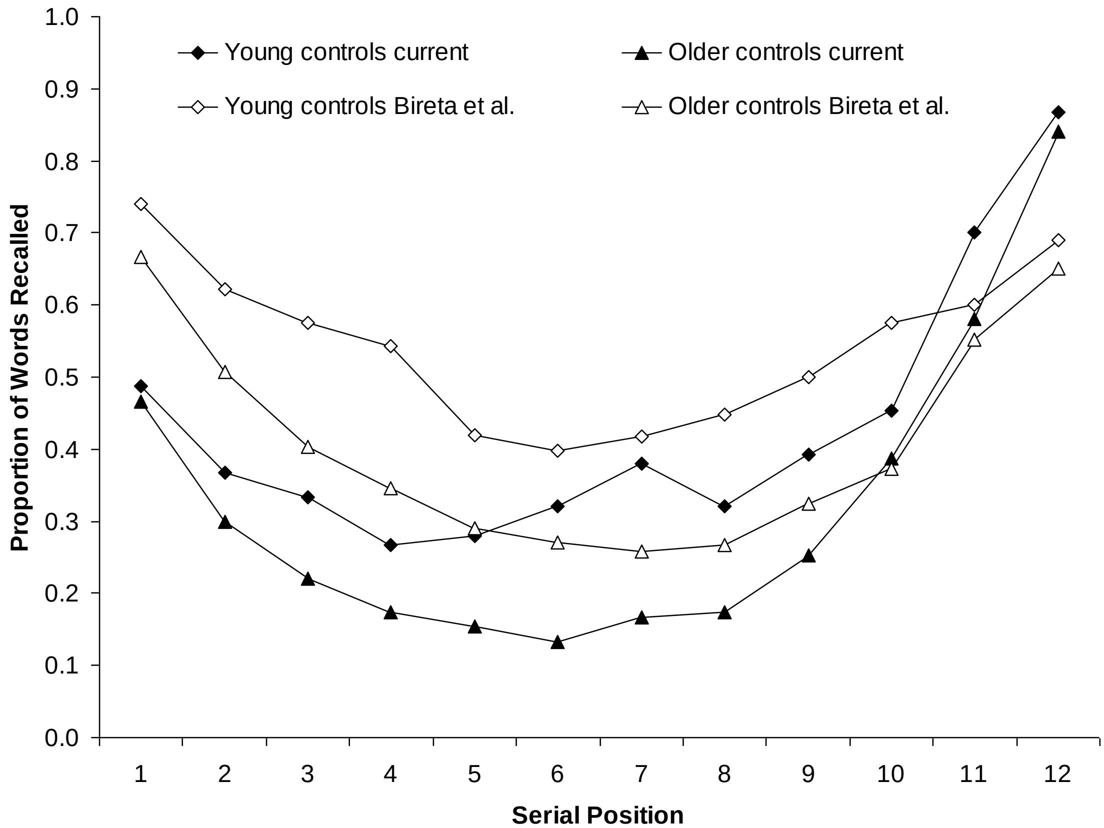

Figure 6. Mean proportion of words recalled as a function of serial position. Data are presented for control lists of Experiment 1 and for fast presentation control lists from Bireta et al. (2008).

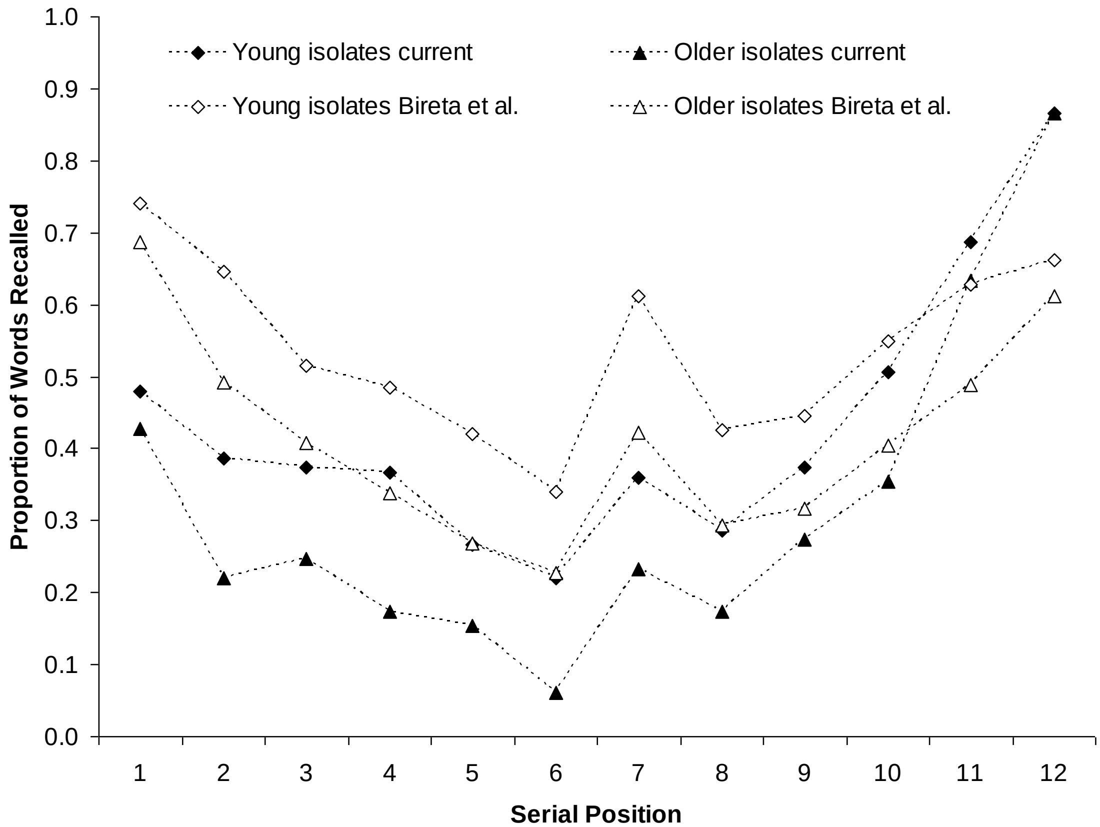

Figure 7. Mean proportion of words recalled as a function of serial position. Data are presented for colour isolation lists of Experiment 1 and for fast presentation colour isolation lists from Bireta et al. (2008).

Awareness. The Bireta et al. (2008) data for probability of recall of the seventh list item were calculated separately for the first and second halves of the experiment and for each participant. As the experiment chose list types randomly for each trial, there was not always an equal amount of controls and isolates in each half; this caused some statistics to be slightly quantitatively (but not qualitatively) different to those reported by Bireta et al. (2008).

To analyse how the isolation effect varied for young and older participants across the experimental period, a 2 (Age: young, older) x 2 (Presentation rate: fast, slow) x 2 (List type: control, isolate) x 2 (Experimental half: first, second) ANOVA was conducted on the seventh position recall data. There was a main effect of age, $F(1, 156)$

= 84.00, *MSE* = 0.09, $p < .001$, with young participants recalling more words on average than older participants. There was a main effect of presentation rate, $F(1, 156) = 13.95$, *MSE* = 0.09, $p < .001$, as more words were recalled for slower presentation rates. A clear isolation effect was present as there was a main effect of list type, $F(1, 156) = 70.80$, *MSE* = 0.06, $p < .001$, with more isolates recalled than controls. Finally, there was also a main effect of experimental half, $F(1, 156) = 5.02$, *MSE* = 0.05, $p < .05$, with more words recalled in the first half. None of the interactions was significant except list type x age, $F(1, 156) = 10.49$, *MSE* = 0.06, $p < .01$, where young participants showed a larger isolation effect than older participants. Crucially, there was no interaction between age, list type and list half ($F < 1$) which could have indicated that increased familiarity/awareness in the second half of the experiment affected young and older participants differently.

The isolation effect strength was calculated by taking the probability of recalling a seventh position control word from the probability of recalling an isolate. Isolation effect strengths were plotted for young and older participants, fast and slow list presentation rates and for first and second experimental list halves (Figure 8). All isolation effects were significant except for older participants in the first half of the experiment during the fast list presentation ($t(39) = 1.53$, *ns*) and for the older participants in the second half of the experiment during the slow list presentation ($t(39) = 1.52$, *ns*). There were no significant differences across experimental halves for the isolation effect strength in each of the four possible categories of data: young participants and fast presentation, young and slow, older and fast, older and slow (all *t*s < 1).

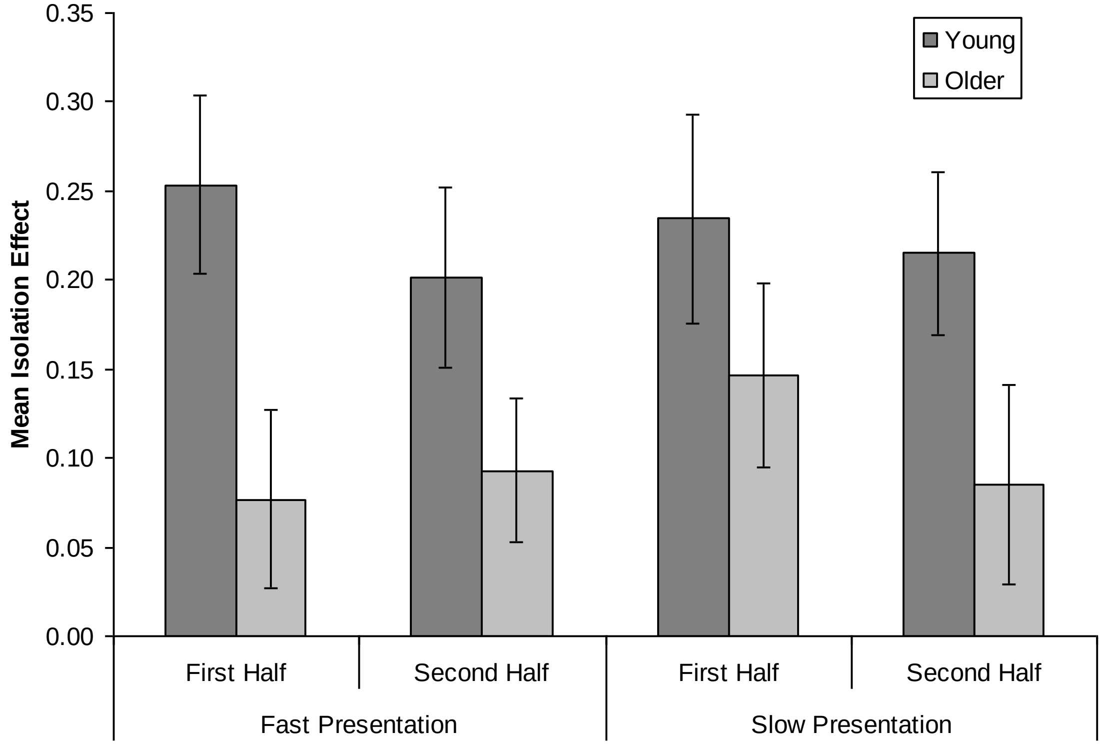

Figure 8. Mean isolation effect for young and older participants from Bireta et al.'s (2008) data. Bars show isolation effect size for fast and slow list presentation, young and older participants and first and second experimental half. Error bars are ±1 *SE*.

Discussion

Comparison between experiments. The comparison between Bireta et al. (2008) and Experiment 1 showed slight differences in recall data alongside the differences in isolation effects. However, the overall patterns in the data were qualitatively very similar. As has been stated in the previous discussion, Experiment 1 produced no significant colour isolation effects whereas Bireta et al. (2008) produced isolation effects that were larger for young participants than for older participants. The main difference between the two experiments is that Experiment 1 had the colour isolation condition mixed in with position and modality isolates as well as controls whereas Bireta et al. (2008) had only colour isolates and controls in the same experimental period. This resulted in 25% of Experiment 1 trials and 50% of Bireta et al. (2008) experiment trials having no isolate. Such differences alter the probability of there being an isolate and therefore the type of anticipation participants would experience. The effects of awareness and anticipation upon the isolation effect are discussed below. Ultimately it has been difficult to comprehensively elucidate a single cause of the isolation effect differences between the two experiments.

There was a larger recency effect for Experiment 1compared to Bireta et al. (2008). This is most likely due to the recall task; in Experiment 1, recall was verbal so participants could say the words they remembered fairly fast, whereas Bireta et al. (2008) used written recall. The faster verbal recall would have helped participants to communicate more words before their memory traces faded from recent memory. Finally, the overall recall levels were greater for Bireta et al. (2008) than for the Experiment 1. It is not clear why this was the case, but the recall differences are most likely to be due to the stimuli used; Bireta et al. (2008) used words with high imageability and familiarity, which could have made them easier to recall.

The isolation effect and awareness. In Experiment 1 there was a marginally significant increase in the magnitude of the isolation effect between earlier and later trials. However, the isolation effect was not significantly different between the first and second halves of Bireta et al. (2008). This indicates that familiarity with isolation list structures does not increase or decrease the strength of the isolation effect. It is reasonable to assume that awareness of isolates could increase towards the end of the experiment and to therefore conclude that awareness does not change the isolation effect. However, the colour isolation used by Bireta et al. (2008) was very easy to perceive and participants may have been aware of the isolates from the onset of the experiment.

In the literature, there is one study that also measured isolation effects across experimental trials (Kelley & Nairne, 2001). They used several experiments to look at how isolation of a stimulus affects its order retention compared to controls. In Kelley and Nairne's (2001) Experiments 1 and 2, words were isolated in a list on the basis of their size at presentation. There was a clear isolation effect with more isolates being correctly repositioned in a list by participants than corresponding control words. There was no significant difference in the isolation effect between trials earlier and later in the experiment. In their Experiment 3, Kelley and Nairne (2001) used another type of isolation based on word generation. Control words were presented normally and isolates were presented with a missing letter (e.g., ph_ne instead of phone). Once again, a clear isolation effect was found but its magnitude was not significantly different between trials earlier and later in the experiment.

The findings of Kelley and Nairne (2001) are congruent with the current analysis of the Bireta et al. (2008) data. Kelley and Nairne (2001) reliably found no isolation effect changes across experimental halves for a range of experiments. The isolation effect does not seem to be altered by increased exposure to isolation and control trials. Elsewhere in the literature, only two experiments were found that directly attempted to manipulate awareness of the isolates in participants (Detterman, 1975; Green, 1958).

Detterman (1975) used volume of sound to generate isolation effects. Lists of 15 words were presented auditorily via a tape recorder at 'normal conversational levels' (p. 614). The isolated word was created by being played back at a much louder level.

Twenty lists were presented to participants with 10 lists containing isolates. In Detterman's Experiment 1, following each list participants wrote down words they could remember in free recall. In Detterman's Experiment 2, following each list participants conducted a yes/no recognition task where nine words were from the list and nine were new words. For both Detterman's Experiments 1 and 2, half of the participants were made explicitly aware that the eighth item would be presented louder on half of the lists. Detterman's Experiment 1 showed that participants unaware of the isolate generated a clear isolation effect; the louder words in position eight were recalled significantly better than the eighth word in control lists. Aware participants showed an inverse isolation effect and remembered more eighth position controls than isolates. In Detterman's Experiment 2, the recognition task did not produce any isolation effects in either the aware or unaware participants. Although the instructions described the isolation effect to the aware participants, it is not clear whether or not the unaware participants were aware of the isolates. However, the awareness manipulation did yield isolation effect differences between the aware and unaware groups in Detterman's Experiment 1.

Green (1958) conducted a follow-up experiment to Green (1956), which aimed to address the effect of surprise upon the isolation effect. The Green (1956) experiment was designed to determine whether increased arousal due to the surprise of the distinctive isolate was responsible for the improved chance of later recalling the isolate compared to control stimuli. The experiment involved the sequential presentation of three-letter nonsense syllables (for example, 'GUB', 'HOF') and three digit numbers. Memory sets were used that each contained two isolates – one of the two isolates was a three digit number within a set of nonsense syllables and the other was a nonsense syllable within a set of three digit numbers. Counterbalancing was created between participants to swap the positions and order of the two isolate types around. Ultimately, the experiment aimed to look at recall of the first and the second isolates within the memory sets. Green (1956) hypothesised that the first isolated stimulus would be recalled better than the second (after factoring out effects of serial position and isolate type) as the second isolate would be less surprising/unexpected. Indeed, this is what the experiment found: The first isolates in the memory sets were significantly more likely to be recalled than the second. It is reasonable to make associations between surprise and attention and Green (1956) argued that the surprise caused by the unexpected isolate raises attention and therefore enriches encoding of the isolate. This could be similar to any effects that awareness might have on an isolate, as one would need to be aware of it to focus attention towards it.

In a follow up experiment, Green (1958) used the same experimental design but manipulated participants' knowledge prior to list presentation. Awareness of the isolate was manipulated by explicitly informing one group of participants about the exact structure of the list (types of stimuli, criteria of isolation and the positions of isolates), whilst another group were simply presented with the list under normal memory instructions. Recall of the isolate was different for the knowledgeable group, with knowledgeable participants showing poorer recall than non-knowledgeable participants of the isolate in the first position and improved recall of the isolate in the second position. Presumably, participants were less surprised or aroused by the earlier isolate but were watching out for the later one. Although awareness was directly manipulated by instruction, these findings still do not confirm the degree to which awareness affects the isolation effect. There was no way of determining how aware the 'ignorant' participants were of the isolate upon presentation with this experimental design.

The findings surrounding awareness of the isolate offer mixed results. This is in part due to the difficulty of ascertaining an awareness measure. Paradoxically, evidence from Detterman (1975) and Green (1958) suggests that awareness of isolates reduces the isolation effect. The reliability of such findings is questionable and in reality other factors such as set size, isolate type, and memory test seem to be much more influential on the isolation effect (Wallace, 1965). In contrast, in the introduction to this chapter it was shown that the general finding of the ageing studies was that participants aware of isolates produce an isolation effect and those unaware do not. This was particularly evident in Geraci et al. (2009) where, within the same experiment, aware participants produced an isolation effect and unaware did not. Both young and older participants showed the same effect and both age groups had around half of participants aware of isolates.

Summary. In the context of ageing research, there was evidence to suggest that reduced awareness of isolates in older participants causes reduced isolation effects (e.g., Geraci et al., 2009; R. E. Smith, 2011). This conclusion is not supported by the current analysis of the Bireta et al. (2008) data. Young and older participants reported equal awareness of the isolate at the end of the test and showed consistent isolation effects throughout the experimental period. Despite these similarities, there was a consistent isolation effect strength difference between young and older participants. This indicates that as participants become more familiar with/aware of experimental memory sets across the experimental period, the isolation effect is not altered. This is a finding that was reliably tested by Kelley and Nairne (2001). Furthermore, there is some evidence to suggest that awareness of the isolates reduces the isolation effect (Detterman, 1975; Green, 1958). These findings are in contradiction to the hypotheses that older participants' isolation effect size reduction/elimination is due to reduced awareness. The results of R. E. Smith (2011) indicated that awareness of the isolate needs to occur when the isolate is being encoded and this offers some insight into discrepancies over the effect of age on the isolation effect across the literature. In conclusion, the literature and current findings do not clarify the role of awareness in the isolation effect, which appears to have different effects for different experimental designs. It may be possible that awareness of the isolates provides older adults with environmental support when associating the isolate to its isolating factor (therefore allowing them to show isolation effects) although this has been shown to be a difficult hypothesis to test. The final section of this chapter forces the isolation effect to be based on associative memory in order to clarify the impact of age-related associative deficits on the isolation effect.

Experiment 2

In order to test if associative deficits may cause differences in the isolation effect, the current experiment aimed to produce an isolating *factor* that was based on associative memory. Theoretically this would mean that older participants, with reduced associative memory, would experience a less isolated isolate than young participants.

The current study was loosely based around Erickson (1963) where the isolating factor was the relationship between items forming pairs in a list. Young participants were shown lists of nine pairs of items; items were either three consonants or three digits (see Table 4 for example list). A given pair could therefore consist of different item types (three letters and three digits), or same item types (three digits and three digits/three letters and three letters). These combinations were used to form isolating criteria based on the relation between the two items in each pair. In a given list of nine pairs, eight would be in one format (e.g., different items - three letters and three digits,) and an isolated pair would be in another format (e.g., same items - three letters and three letters). The isolated pair was always in the fifth position of a list. Participants completed cued recall where they were shown the left item of each pair and had to recall the corresponding right item. It was found that isolated items were recalled more successfully than non-isolated items. This demonstrated that an associative relation could form an isolating factor.

Table 4

Example Experimental List from Erickson (1963)

List Pairs
SWJ-217
BJN-821
RKD-764
CTG-472
KSC-ZNH
086-DXR
590-GDP
305-JPZ
953-LFS

The current study created lists where the relationship between pairs of items was the isolating factor. If older adults form weaker associations between pairs of items than do young adults, then this experimental set up should produce a smaller isolation effect in older adults compared to young adults. This is because if older adults form weaker associations between items then the isolate would appear less distinct for them compared to young participants. It was discussed earlier in this chapter that distinctiveness of the isolate compared to non isolates is a key factor for creation of an isolation effect: In his review of the isolation effect, Wallace (1965) concluded that the magnitude of the isolation effect varied directly with the degree of isolation.

Method

Participants. Twenty-four young adults (8 female), aged 18-24 years (M = 20.4, SD = 1.38), and 20 healthy older adults (11 female), aged 61-77 years (M = 68.7, SD = 4.3), took part in the experiment.[8] Participants were recruited on a voluntary basis and were offered no financial incentives for taking part. The mean number of years of education was obtained for each participant: Young participants had completed significantly more years of education than older participants, $t(42) = 8.77$, $p < .001$ (young M = 17.8, SD = 0.7; older M = 15.7, SD = 0.8).

To assess cognitive functioning, participants completed the Digit Symbol Substitution task (Wechsler, 1981) as a measure of perceptual-motor or processing speed. They also completed the multiple choice part of the Mill Hill vocabulary test (Raven et al., 1988) as a measure of crystallised intelligence. The results were consistent with the literature (e.g., Horn & Cattell, 1967; Salthouse, 1991). Young participants were significantly faster at the digit symbol substitution task than older participants, $t(42) = 6.13$, $p < .001$ (young M = 60.9, SD = 13.4; older M = 40.3, SD = 7.4). For the vocabulary test, young participants scored significantly lower than older participants, $t(42) = 2.19$, $p < .05$ (young M = 16.8, SD = 3.6; older M = 19.4, SD = 4.2).

[8] Data were gathered by Lauren Brawn and Charlotte Gillingham.

Materials. Stimuli were combinations of four letters/digits separated by a hyphen in the middle (e.g., BC-12). All of the letters in the alphabet were used excluding vowels and the letter ‘o’. All of the digits were used except zero. Vowels were excluded to prevent the formation of syllables which could enhance memory by chunking (Miller, 1956). The letter ‘o’ and digit ‘0’were excluded because of their similarity in appearance.

For a given trial, participants viewed five sets of stimuli[9] sequentially at a rate of 5 s per set (see Figure 9). Following this there was a 5-s delay, then a recognition test. There were eight control trials. The stimuli in control trials always had two letters on the left of the hyphen and two digits on the right and vice versa. There were eight isolation trials where the third set of stimuli was presented as an isolate based on the arrangement of letters and digits: There were four number-isolation trials, the five sets of stimuli being arranged the same as control trials apart from the third set, which consisted of entirely digits both on the left and right of the hyphen (e.g., 12-34). There were also four letter-isolation trials, with the five sets of stimuli arranged the same as controls apart from the third set, which consisted of letters on both sides of the hyphen (e.g., BC-DF). At all stages, the letters and digits were randomly selected under the constraint that no specific pair of letters/digits could appear twice within a trial and no two letters/digits were the same within a given memory stimuli (e.g., not BB-22).

The recognition test after each trial always consisted of pairs of letters/digits from the trial (see Figure 9). Therefore, participants needed to be aware of the associations between the left and right pairs in order to respond correctly. Participants

---

[9] A small pilot study was conducted with four sets and five sets of stimuli with young participants. Five sets of stimuli was chosen as participants were not performing at floor or ceiling and the odd number of stimuli allowed an isolate to be placed exactly in the centre of the list (position three).

completed five recognition tests in each trial – one for each set of memory stimuli. An individual recognition test consisted of a pair of letters/digits on the left of the screen which was originally presented on the left in one of the memory sets and a forced choice of two pairs of letters/digits on the right (which were both originally presented on the right of two memory sets). One of the right-hand pairs would correspond to the pair that was originally shown with the left-hand pair during the study period. Both of the right-hand pairs were always of the same type (i.e., both digit pairs or both letter pairs) so that the pair type did not indicate the correct answer. The five recognition tests covered each of the five left-hand pairs of stimuli and the left-hand test pairs were randomly selected so that participants were not tested on the stimuli in the same order they were presented.

The letters/digits were displayed on a laptop computer screen at a viewing distance of approximately 60 cm and the height of the letters/digits corresponded to approximately 1.5° viewing angle.

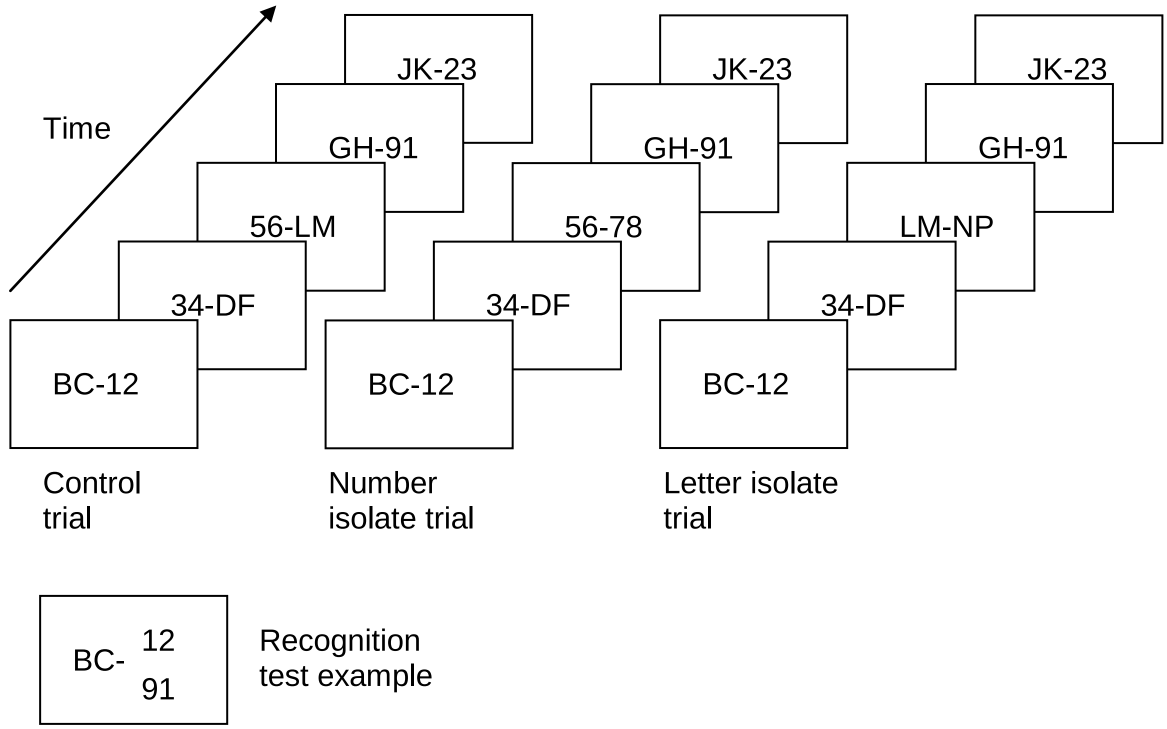

Figure 9. Top: Examples of control, number isolate and letter isolate trials. Bottom: Example of the 2 alternative forced choice recognition test.

Procedure. Participants were informed that they would be shown five sets of memory stimuli and their memory for each of the associations would be tested later. After each memory set, they completed five recognition tests. They had to complete a forced choice to identify which of the two right-hand options (top or bottom) was originally shown with the left-hand pair of letters/digits (see Figure 9 for example). For the recognition tests, participants were instructed to press the 't' key on a keyboard if they thought the top option was originally displayed with the left pair or the 'b' key if they thought the bottom option was originally displayed with the left pair. Participants were allowed as long as they needed to respond. This procedure was repeated for each of the 16 trials. At the end of the experiment participants were asked if they noticed anything unusual about the stimuli in order to establish awareness of the isolate.

Results

To begin with, the two isolation conditions were compared to assess any differences between number and letter isolates. A 2 (Age: young, older) x 2 (Isolation type: numbers, letters) x 5 (Serial position 1-5) repeated measures ANOVA was conducted on the proportion of correctly recognised letter/digit associations. There was no main effect of isolation type, $F(1, 42) = 1.13$, $MSE = 0.07$, *ns*, with letter and number isolates producing the same level of recognition performance. There were also no significant interactions between any of the factors, which indicated that the isolation type was not differentially affecting recognition performance for the different age groups and serial positions. Therefore, the following analysis was conducted using the average of number and letter isolation conditions.

A 2 (Age: young, older) x 2 (Condition: control list, isolation list) x 5 (Serial position 1-5) repeated measures ANOVA was conducted on the proportion of correctly recognised letter/digit associations (see Figure 10 for means). There was a main effect of age, $F(1, 42) = 5.97$, $MSE = 0.77$, $p < .05$, with young participants performing better than older participants on average (*M* (*SD*) = 0.71 (0.24) and 0.62 (0.23), respectively). There was no main effect of condition, $F(1, 42) = 1.75$, $MSE = 0.05$, *ns*, with overall memory for letter/digit associations equivalent in control and isolate conditions. There was a main effect of serial position, $F(4, 168) = 3.07$, $MSE = 0.08$, $p < .05$, with differing recall across serial position due to the isolates and primacy/recency effects. The crucial interaction between serial position and condition was significant, $F(4, 168) = 3.13$, $MSE = 0.10$, $p < .05$, indicating the presence of an isolation effect as the participants were differentially remembering pairs across serial positions for control and isolation lists. None of the other interactions was significant.

In order to directly establish if the experiment produced an isolation effect, a 2 (Age: young, older) x 2 (Condition: control list, isolation list) repeated measures ANOVA was conducted on just the data for the third serial position. There was no significant difference between young and older adults, $F(1\ 42) = 2.48$, $MSE = 0.10$, *ns*, (*M* (*SD*)= 0.76 (0.24) and 0.68 (0.24), respectively). There was a main effect of condition, $F(1, 42) = 13.73$, $MSE = 0.03$, $p < .001$, with isolates being recalled significantly better than non isolates on average (*M* (*SD*) = 0.77 (0.25) and 0.63 (0.18), respectively). This demonstrates that an isolation effect was present in the data. Importantly, with regard to the hypothesis, there was no interaction between condition and age ($F < 1$). Therefore, young and older adults did not significantly differ in the magnitude of isolation effects produced.

Finally, young and older participants' third serial position data were analysed separately in paired sample *t*-tests (third position control vs. third position isolate) to confirm that both age groups produced isolation effects. For both young and older participants, there was a significant isolation effect: $t(23) = 2.93$, $p < .01$, and $t(19) = 2.39$, $p < .05$, respectively. The difference between mean control and isolate list third position recognition scores was 0.16 for young and 0.11 for older participants. This suggests that young adults produced a numerically larger isolation effect than older adults. It can be seen in Figure 10, however, that this difference is partly driven by a reduction in third position recognition in young participants' control lists.

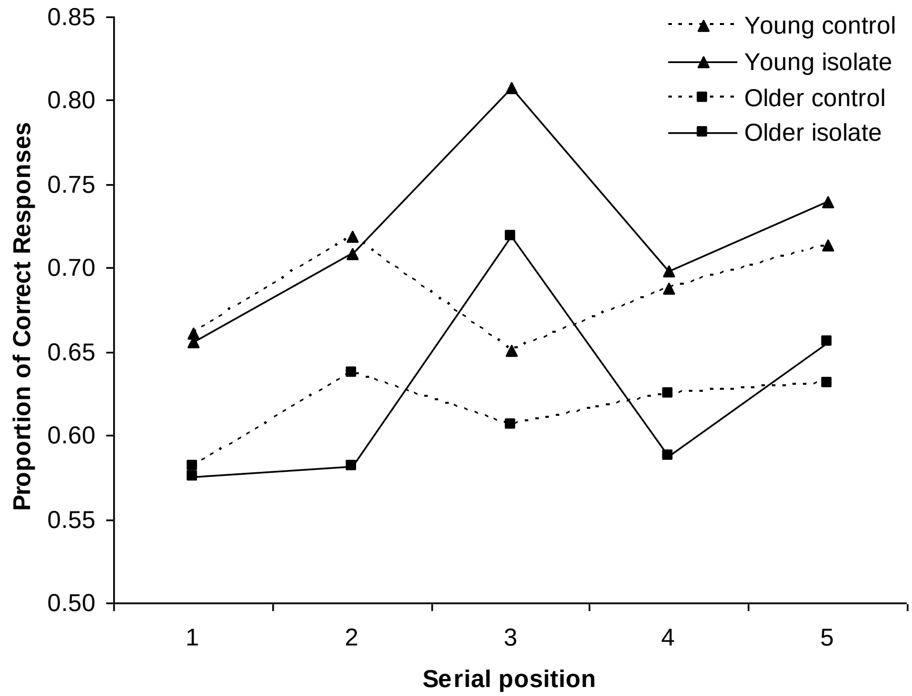

Figure 10. Mean proportion of correct responses as a function of serial position for control and isolate conditions and for young and older adults.

In order to establish further any age differences in the isolation effect, a separate analysis was completed by comparing the average of all non-isolate serial positions (Serial positions 1, 2, 4 and 5 averaged) to isolate positions 3 for both control and isolate lists. This provided a measure of the isolation effect which used all of the experimental data. A 2 (Age: young, older) x 2 (Isolate position: positions 1, 2, 4, 5 averaged, position 3) x 2 (Condition: control, isolate) repeated measures ANOVA was conducted on the proportion of correctly recognised letter/digit associations. There was a main effect of age, $F(1, 42) = 4.88$, $MSE = 0.05$, $p < .05$, with young participants performing better than older participants (M (SD) = 0.71 (0.24) and 0.62 (0.23), respectively). There was a main effect of condition, $F(1, 24) = 9.33$, $MSE = 0.02$, $p < .001$, with control list associations being recognised worse than isolate list associations (*M* (*SD*) = 0.64 (0.13) and 0.70 (0.14), respectively). This was driven by the increased weighting of the isolate due to averaging of non isolated positions together. There was also a main effect of position, $F(1, 42) = 5.96$, $MSE = 0.01$, $p < .05$, with position three associations being recognised better than positions 1, 2, 4 and 5 averaged (*M* (*SD*) = 0.70 (0.14) and 0.66 (0.12), respectively). This effect was also driven by the isolate. An isolation effect was apparent as there was an interaction between condition and isolate position, $F(1, 42) = 13.51$, $MSE = 0.02$, $p = .001$. There was a much larger recognition performance difference between control and isolate lists for third position associations than for the average of first, second, fourth and fifth position associations, with higher recognition for the isolated associations. This measure yielded an isolation effect of 0.15 for young participants and 0.13 for older participants.[10] Crucially, as with the earlier analysis, there was no evidence of different isolation effects between young and older participants: There was no triple interaction between condition, isolation position and age ($F < 1$). No other interactions were significant (all *F*s < 1).

Discussion

The experiment successfully produced an isolation effect in both young and older adults. Despite the fact that the isolating factor was based on associative memory, there was no significant difference in the magnitude of the isolation effect between young and older participants. Older adults performed worse than young participants

---

[10] A measure of isolation effects using all of the data was calculated by calculating the difference between the third position performance and the non-third position performance (positions 1, 2, 4, 5 averaged vs. position 3) for control and isolate lists. Then this difference for isolation lists was subtracted from this difference for control lists. That is, (((control 1+ 2 + 4 + 5)/4) - control 3) – (((isolate 1 + 2 + 4 + 5)/4) – isolate 3).

overall; this demonstrates that the associations they formed between the pairs of items in the memory test were weaker than those formed by young participants. Even though older adults formed weaker associations, this did not alter the isolation effect evident in their responses. This demonstrates that deficits in associative memory are not strongly linked to the magnitude of the isolation effect. Therefore, the proposal by Bireta et al. (2008) that associative deficits in older adults can reduce the magnitude of the isolation effect is not supported by the current data.

Previous studies have shown with young participants that the learning of associations between pairs of items is facilitated when a given pair is isolated (Erickson, 1963; Kimble & Dufort, 1955; Nachmias, Gleitman, & McKenna, 1961). It is therefore feasible that as older adults' associative memory performance was enhanced by isolation, the isolating factor was also enhanced and hence no age differences were observed. This highlights the unusual circular nature of the current experiment whereby the isolating factor (an associative memory) is actually enhanced by increased memory performance due to isolation. Thus there is evidence here that the isolation effect/distinctiveness of stimuli can reduce associative deficits in older adults. Both young and older adults reported unanimous awareness of the isolate. This is congruent with the idea that differential awareness between young and older adults contributes to isolation effect differences across age. This would suggest that there were no differences in the isolation effect across age because there were also no differences in awareness.

General Discussion

There remain mixed results surrounding age differences in the isolation effect and its mediation by an age-related associative deficit. The associative deficit hypothesis suggests that older participants would form a weaker bond than young participants between a stimulus and its isolating factor. For older participants, associative memory deficits may therefore reduce the degree of isolation, limiting the benefit from enhanced encoding or retrieval compared to that found in young participants. This provides a viable explanation of why some studies show a reduced isolation effect in older participants compared to young participants. On the other hand some experiments (including the current Experiments 1 and 2) show no age differences in the isolation effect. As age-related associative deficits are a widely robust finding, it would be expected to see age-related isolation effect differences all of the time. Since this is not the case, it may be that associative deficits are not responsible for age differences in the isolation effect. Alternatively, the associative deficit may be alleviated by isolation, which provides environmental support at encoding to the processing of isolates (cf. Craik, 1986).

A possible hypothesis was that age differences in the isolation effect are linked to a differential level of awareness of the isolate in young and older participants. However, this hypothesis was shown not to be responsible for the age differences in the study by Bireta et al. (2008). Awareness of the isolate is a difficult phenomenon to measure and does not seem to strongly influence the isolation effect. Exploring this area further in the context of age differences is unlikely to address the role of age-related associative deficits in the isolation effect.

Age-related associative deficits and age-related awareness differences were considered as causes for age differences in the isolation effect. Despite a detailed analysis of these areas, the mixed results in the literature surrounding age differences in the isolation effect remain unclear. There is no obvious agreement in the literature that there are age differences in the isolation effect. It has been shown here that associative deficits are not necessarily linked to the isolation effect. Thus, using isolation effects to develop an understanding of the associative deficit hypothesis is currently not a viable avenue of research. With the current understanding, further exploration into age-related differences in the isolation effect would not be hypothesis driven. Therefore, the following chapters explore the associative deficit hypothesis by other means. The next chapter continues the theme of distinctiveness and associative memory in an applied study that looks at eye witness identification of distinctive faces.

Chapter 4: Associative Deficits and Identifying Faces with Distinctive Features

Age-related associative memory deficits are not only observed between highly controlled experimental stimuli such as word pairs (e.g., Naveh-Benjamin, 2000) but also for more rich stimuli such as pairs of pictures (e.g., Naveh-Benjamin, Hussain et al., 2003) and pairs of faces (e.g., Bastin & Van der Linden, 2006; Rhodes, Castel, & Jacoby, 2008). Even when associative memory tasks mimic everyday uses of memory such as associating a name to a face, older adults show a reduced performance compared to young adults after taking into account memory for the faces and names individually (e.g., Naveh-Benjamin, Guez, Kilb et al., 2004; Naveh-Benjamin et al., 2009). Along similar lines, older adults have also shown memory deficits relative to young adults for associating a person to an action they were completing (Old & Naveh-Benjamin, 2008b). The current study addresses age-related associative deficits in a practical context, that is, in the recognition of faces in a police lineup. In addition, the current study aims to further investigate how congruency between study and test stimuli mediates associative deficits.

In a recent study, Zarkadi, Wade and Stewart (2009) addressed an important issue related to identifying culprits of crimes in police lineups who have distinctive features (e.g., a moustache). They investigated the most suitable method to display suspects with distinctive features among other lineup members so that they would not be identified purely on the basis of possessing the distinctive feature. This is a surprisingly common problem in police investigations and around one third of all lineups in England

and Wales need to be digitally manipulated[11] to avoid distinctive suspects from standing out (see Zarkadi et al., 2009). Also, in a survey of US police officers' practices for lineup preparation and conduct, 70% of officers reported using methods to avoid a suspect with distinctive features from standing out (Wogalter et al., 2004). In the context of ageing research, age-related associative deficits may influence the association in memory between a person and a distinctive feature that they possess. This may therefore impact on the most suitable method of presenting lineups to older adults.

When creating lineups for criminal identification, police often have details about the culprit's appearance provided by witnesses. A problem may occur with lineups if a suspect is reported to have a distinctive feature. This is because they may stand out in a lineup and they may be easier to identify, thus making the procedure unreliable for criminal conviction. Furthermore, if an innocent person with the same reported distinctive feature is placed in a lineup among people who do not have that feature, then the innocent person is likely to be incorrectly identified as a criminal (Wells, Rydell, & Seelau, 1993; Wells et al., 1998). This is especially true for simultaneous lineups (which account for 90% of all police lineups; Wogalter et al., 2004) where witnesses are more likely to use a relative judgement strategy (Wells et al., 1998). That is, witnesses express a tendency to choose the lineup member who looks the *most* like the culprit they remember, even if the selected lineup member is not actually the same person. To address solutions to this problem, Zarkadi et al. (2009) investigated the most suitable method for digitally manipulating faces in photographic lineups to stop suspects with distinctive features from standing out. Digital manipulation of images is the most

[11] The majority of modern lineups are created with photo arrays, not lines of people (Wogalter, Malpass, & McQuiston, 2004).

practical approach for police officers to use because it is much less expensive than finding foils who have the same feature. Indeed, it may be impossible to find foils with highly specific features like a facial tattoo.

Two methods commonly used by police officers were compared by Zarkadi et al. (2009): *replication*, where a distinctive feature in a culprit was digitally added to all of the foils in a lineup, and *concealment*, where the distinctive feature was removed from the culprit (the target) and the target appeared among foils with non-distinctive faces. There is currently no set procedure in the UK or the US for which method to use with real suspects and the decision is made by individual police officers (Zarkadi et al., 2009). To conduct the test, Zarkadi et al. initially showed participants a memory set of 32 faces which contained 6 target faces with distinctive features (e.g., a tattoo). After a short delay they then presented lineups to participants which each showed 6 faces simultaneously. Half of the lineups used the replication method to uniformly present lineup members and half used the concealment method. They found that replication was a more successful technique – it resulted in more target identifications than concealment in target present lineups and it did not result in increased foil identification in target absent lineups. This finding is in line with the encoding specificity hypothesis (Tulving & Thompson, 1973) where memory performance is improved when encoding and retrieval occur in similar contexts. This is because for replication lineups, the target appears exactly as it did during encoding, but for concealment it does not. Zarkadi et al. (2009) also found that the hybrid similarity model (Nosofsky & Zaki, 2003) predicted the pattern of results, whereas standard global familiarity models that do not take into account the effects of distinctive features did not (e.g., Valentine & Ferrara, 1991).

Unlike global familiarity models, which predict similar performance under replication and concealment for target present lineups, the hybrid similarity model takes distinctiveness into account and is therefore able to predict the findings of Zarkadi et al. (2009). Both the global familiarity model and the hybrid similarity model make use of summed similarity, where items in recognition tests (e.g., lineups) are compared to *all* items in memory and if the summed similarity evoked (i.e., overall familiarity to all items) crosses a certain threshold, the item is recognised. Both models also predict that target and foil faces will evoke more familiarity in replication lineups than concealment lineups because for replication lineups they possess a distinctive feature that is shared with a study face. However, the hybrid similarity model applies a multiplicative boost to similarity when faces at test share distinctive features with faces in memory. This means that for replication lineups, when target and foil faces share a distinctive feature with a given study face, similarity is boosted more for target faces than for foils (because target faces initially evoke more similarity than foils and thus there is more similarity to boost). This means that the absolute similarity evoked by targets is larger than foils for the hybrid similarity model so performance is higher as they appear more discriminable from foils.

The effect of age on recognition of faces with distinctive features may alter the benefit of replication over concealment. It is well established in the literature that older adults generally have poorer memory than young adults (e.g., Zacks et al., 2000) and this finding also occurs for memory of faces (e.g., Bartlett, Leslie, Tubbs, & Fulton, 1989; Grady et al., 1995; Naveh-Benjamin, Guez, Kilb et al., 2004). Memon, Gabbert and Hope (2004) found that older adults were more prone to selecting foils in target absent lineups than young adults. Similarly, in a recognition memory test, Bartlett et al. (1989) found that older adults performed worse than young adults because they identified more foils at test. Interestingly, in Bartlett et al.'s (1989) Experiment 1, older adults were just as good as young adults at recognising seen before but altered faces (e.g., a face with a change of expression between study and test). However, in Experiment 2 they could not explicitly define what had been altered on a face as well as young adults. This indicated that older adults may behave differently to young adults when characteristics of a face are changed between study and test, possibly because they form weaker associations between a face and how it is presented.

If older adults express associative deficits when forming links between faces and their distinctive features, then there may be less of an effect of distinctive features in lineup recognition tests. During encoding, older adults may not have sufficient cognitive resources to encode distinctive features at the same time as the faces. Smith (2011) and Geraci and Rajaram (2002) argued that the processing of similarity and difference may require cognitive resources and Smith pointed out that this may cause older adults (who show reduced cognitive resources, e.g., Craik, 1982) to have difficulty processing distinctiveness. Therefore, if older adults have weak links between faces and their distinctive features the effect of distinctive features will be reduced and so will the difference between replication and concealment lineups memory performance.

On the other hand, age-related associative deficits may have little impact on the memory performance difference between replication and concealment lineups and older adults may show a similar benefit for replication over concealment lineups as is seen in young adults. The age-related deficits in associative recognition tests are often driven by increased false alarms to lures whilst endorsement of seen-before associations remains relatively intact (e.g., Castel & Craik, 2003; Healy, Light, & Chung, 2005). The results from Zarkadi et al. (2009) showed that there was little difference between replication and concealment on target absent trials. This could mean that age-related associative deficits may have little effect on target-present trials. If older adults show similar endorsement performance to young adults on associative memory tests, then they may also benefit equally to young adults from replication lineups compared to concealment lineups.

The current study replicated Zarkadi et al. (2009) with young and older adults under more favourable encoding conditions (50% longer study time to improve memory performance). This allowed the two age groups to be compared without the older adults performing at floor. The current study therefore extended Zarkadi et al.'s (2009) research to establish if replication was also more beneficial to memory compared to concealment with older witnesses. From an applied point of view, this would be useful to determine if replication should be recommended to police officers conducting lineups for witnesses of all ages. Furthermore, the current study aimed to explore the effect of distinctiveness in relation to age-related associative deficits by modelling young and older adults' data with the hybrid similarity model. This would determine if older adults make use of the presence and/or absence of distinctive features to the same extent as young adults during recognition memory tests.

Experiment 3

Method

Participants. Sixty young adults (30 female) aged 18-24 years (*M* = 20.4, *SD* = 1.4), and 90 older adults (51 female) aged 61-91 years (*M* = 74.2, *SD* = 7.4), took part in the experiment.[12] Young participants were an opportunity sample. Some of the older participants were recruited from the University of Warwick Age and Memory Study volunteer panel that was populated by local advertisements and from the local community. Other older adults were recruited through friends, family and community groups. All older adults were living independently. Participants were not offered any financial incentives for participation.

To assess cognitive functioning, participants completed the Digit Symbol Substitution task (Wechsler, 1981) as a measure of perceptual-motor processing speed. They also completed the multiple choice part of the Mill Hill vocabulary test (Raven et al., 1988) as a measure of crystallised intelligence. The results were consistent with the literature (e.g., Horn & Cattell, 1967; Salthouse, 1991). Young participants were significantly faster than older participants at the Digit Symbol Substitution Task, $t(148) = 12.21$, $p < .001$ (young *M* = 67.0, *SD* = 12.5; older *M* = 42.5, *SD* = 11.6). For the vocabulary test, young participants scored significantly lower than older participants, $t(148) = 3.79$, $p < .001$ (young *M* = 16.5, *SD* = 4.4; older *M* = 19.7, *SD* = 5.5).

Materials. The stimuli used consisted of black and white images of 98 faces all taken from those employed by Zarkadi et al. (2009). The images were obtained from Florida's Department of Corrections website – all images were of inmates aged 24 years

[12] Data from the young adults and 71% of the older adults were collected by Hannah Watts and Natalie Woods.

old who had short brown hair and brown eyes. All of the faces had neutral expressions, were looking directly at the camera and were in front of a neutral grey background. All of the images had any distinguishing features removed (by Zarkadi et al., 2009) such as birthmarks or facial hair using Adobe Photoshop CS2. Forty-two of the 98 different faces were manipulated to have a distinctive feature added. There were six different distinctive features (Bruise, Mole, Moustache, Piercing, Scar and Tattoo) each given to seven faces. This produced an extra 42 images with uniform distinctive features (see Figure 11 for examples).

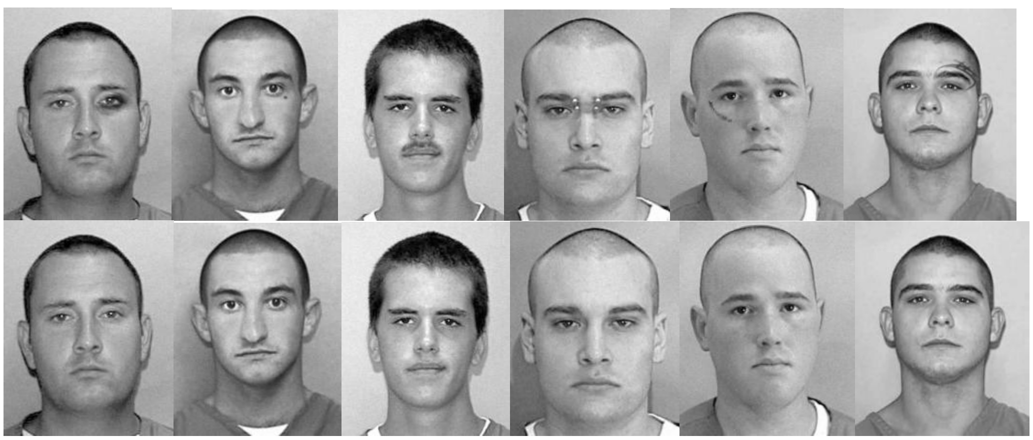

Figure 11. Examples of distinctive features (top) added to plain faces (bottom). From left to right: bruise, mole, moustache, piercing, scar, tattoo.

Design. A recognition memory experiment was conducted with each participant. The memory set consisted of 32 faces; 26 were plain and the remaining six were target faces that each had one of the six distinctive features. To test memory of these faces, four different types of six-face lineups were constructed:

Target-present replication. A distinctive face from the memory set was present and all five foils had the same distinctive feature.

Target-present concealment. A distinctive face from the memory set was present but with that distinctive feature absent and plain faces with no distinctive features were used for the remaining five foils.

Target-absent replication. All six members of the lineups were new faces but all had the same distinctive feature as one of the six distinctive faces from the original memory set.

Target-absent concealment. All of the faces were new and had no distinctive features.

Randomisation. Randomisation was conducted separately for each participant. The six distinctive faces in the memory set were the target faces that would be used in the lineups. In the original memory set, the same 26 non-distinctive faces were always used and appeared in a random order. The six distinctive faces in the study memory set were randomly chosen from the set of 42 distinctive faces under the constraint that each of the six had different distinctive features. The six distinctive faces were placed randomly in the study memory set.

For the 12 lineups themselves, the four different lineup types were tested three times each and could appear in any order (i.e., non-blocked). All replication lineup lure faces were selected randomly from the relevant bank of faces with the required feature. All concealment faces were selected randomly from non-distinctive faces. The positioning of faces in the lineup was random and the target faces could appear in any location.

Procedure. At study, each face was displayed sequentially in the centre of a laptop screen at a rate of 3 s per face (Zarkadi et al., 2009, used 2 s per face). Participants were instructed to remember all of the faces for a later memory test. They were instructed to remember the individuals themselves but were informed that they may appear differently in the following memory test. After the memory set display, participants completed the Digit Symbol Substitution Task for a fixed duration of 90 s before the memory test (note that Zarkadi et al., 2009, used a 5-minute delay).

For the memory test, participants viewed a lineup screen which showed six images of faces arranged in two rows of three faces. They were asked to indicate via a button press of numbers one to six on the laptop keyboard which face they had seen before, or if they recognised no faces to press the number zero. They were informed that they could only respond once and that there would not always be a face from the memory set in the lineup. In total there were 12 lineups and the next lineup appeared immediately after a response was made.

Results

The responses from participants for each of the 12 lineups were categorised into three groups: A *target* response was when they correctly identified a target face from the memory set, a *foil* response was when they incorrectly identified a foil in the lineup, and a *none* response was when they correctly or incorrectly decided that none of the faces in the lineup had been seen before. Figure 12 shows the proportion of responses falling into each category for the four different types of lineup and for young and older adults. To begin with, responses for each response category (*target*, *foil* and *none*) were entered individually into 2 (Age: young, older) x 2 (Lineup Type: replication, concealment) repeated measures ANOVAs separately for target-present and target-absent lineups.

For target-present lineups, *target* responses were higher in young adults compared to older adults, $F(1, 148) = 19.95$, $MSE = 0.11$, $p < .001$. More targets were identified in the replication lineups than in the concealment lineups, $F(1, 148) = 32.85$, $MSE = 0.07$, $p < .001$. There was also an interaction between age and lineup type, $F(1, 148) = 8.74$, $MSE = 0.07$, $p < .01$, with older adults benefiting less from the replication lineups over concealment lineups compared to young adults.

For target-present lineups, *foil* responses were lower in young adults compared to older adults, $F(1, 148) = 16.59$, $MSE = 0.12$, $p < .001$. More foils were identified in the concealment lineups than in the replication lineups, $F(1, 148) = 15.81$, $MSE = 0.06$, $p < .001$, indicating that replication lineups are better than concealment lineups because they reduce false identifications. There was an interaction between age and lineup type, $F(1, 148) = 6.06$, $MSE = 0.06$, $p < .05$, again with older adults benefiting less from the replication lineups compared to young adults.

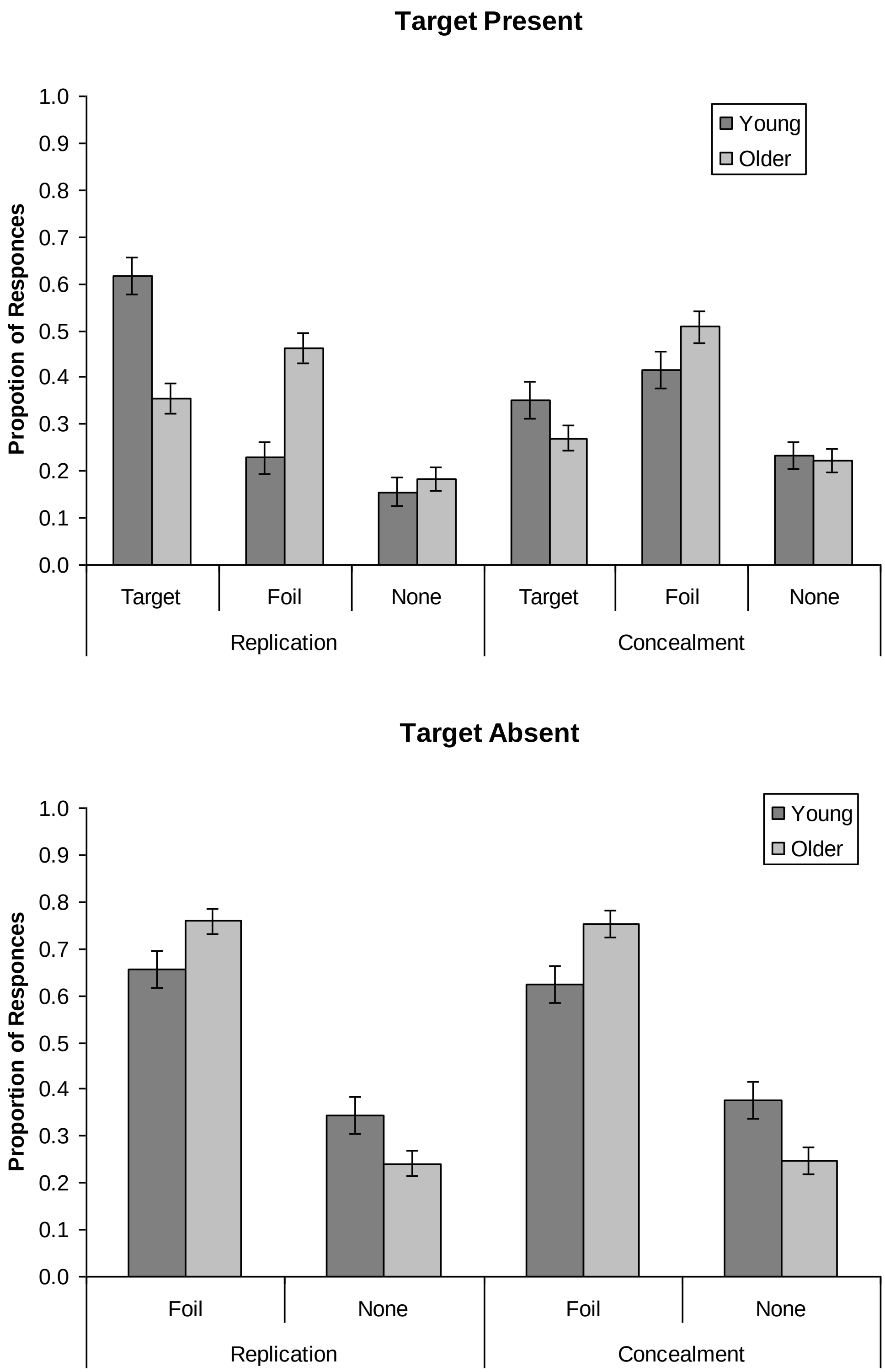

Figure 12. Mean proportion of responses in each response category (identifying a target, foil or none of the faces) in replication and concealment lineups for young and older adults and for target-present (top) and target-absent (bottom) trials. Error bars are ± 1 *SE*.

For target-present lineups, *none* responses were similar in young and older adults, $F < 1$. More *none* selections were made in the concealment than in the replication lineups, $F(1, 148) = 5.33$, $MSE = 0.05$, $p < .05$, again indicating that the replication lineups result in improved detection of seen-before faces. There was no interaction between age and lineup type, $F < 1$.

For target-absent lineups, *foil* responses were lower in young adults compared to older adults, $F(1, 148) = 9.43$, $MSE = 0.10$, $p < .01$. There was no main effect of lineup type or an interaction between age and lineup type, *F*s < 1. The *none* responses were statistically identical as *foil* and *none* response proportions must sum to one. Therefore, young adults made more correct *none* responses than older adults, demonstrating superior performance in young adults.

To summarise, young adults demonstrated better memory performance than older adults in that they detected more targets and endorsed fewer foils. When a target was present, the replication lineups produced superior memory performance compared to the concealment lineups. When a target was not present, the replication and concealment lineups showed similar levels of performance. These data successfully replicate the findings from Zarkadi et al. (2009). Finally, for target present lineups both young and older adults demonstrated superior performance for the replication lineups compared to concealment, but the older adults benefited to a significantly lesser extent.[13]

---

[13] After completing the experiment, a small number of older adults reported confusion during their first encounter of replication lineups. When they encountered a replication lineup for the first time, they reported responding to the first face they looked at because it had a distinctive feature. Only after responding did they realise that all the faces had distinctive features. The main data were reanalysed by excluding the first replication lineup and the first concealment lineup that each participant encountered. The results were qualitatively the same as those found with the whole data set. Young adults still performed better than older adults, replication lineups produced better identification than concealment lineups, and the benefit of replication over concealment was mainly present in young adults.

Target-present lineups were analysed further to ensure that older adults were not performing at chance levels. This was to check that the difference in benefit of replication over concealment across the two age groups was not due to floor effects. Only responses that endorsed one of the six lineup faces were analysed (i.e., *target* and *foil* responses). The chance of a participant identifying a target, given that they made a selection is 1/6 (corresponding to a proportion of .17). The proportion of endorsement responses that were targets was calculated and entered into a 2 (Age: young, older) x 2 (Lineup Type: replication, concealment) repeated measures ANOVA[14] (see Figure 13 for means). The pattern of results was similar to the previous analysis. Young adults identified a larger proportion of targets compared to older adults, $F(1, 146) = 17.65$, $MSE = 0.15$, $p < .001$. A larger proportion of targets were identified for replication lineups compared to concealment lineups, $F(1, 146) = 18.68$, $MSE = 0.10$, $p < .001$. There was also an interaction between age and lineup type, $F(1, 146) = 7.26$, $MSE = 0.10$, $p < .01$. Paired *t*-tests between replication and concealment lineup performance confirmed that young adults benefited from replication, $t(59) = 4.40$, $p < .001$, but older adults did not, $t(87) = 1.31$, *ns*. Additionally, the proportion of target endorsements were above chance (> .17) on replication lineups for young adults, $t(59) = 13.22$, $p < .001$, and older adults, $t(88) = 7.07$, $p < .001$, and also on concealment lineups for young adults, $t(59) = 5.83$, $p < .001$, and older adults, $t(88) = 5.25$, $p < .001$.

[14] Two older adults made no endorsements for one of the lineup types with target-present lineups (one for replication lineups and one for concealment lineups) so were not included in this analysis.

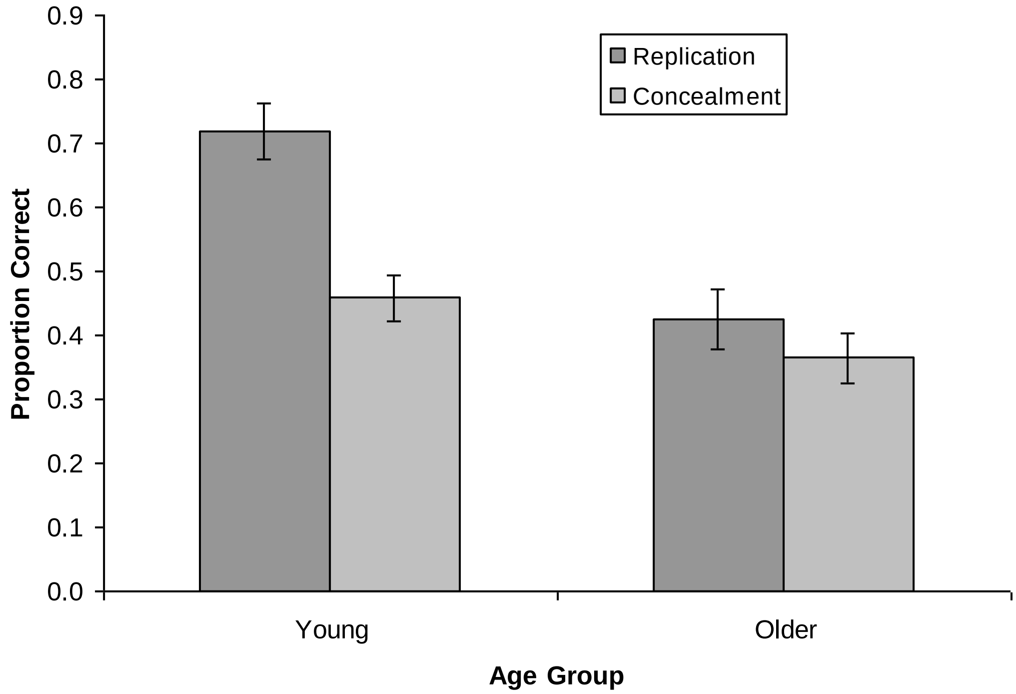

Figure 13. Proportion of endorsements in target-present lineups that were correct for replication and concealment lineups for young and older adults. Error bars are ± 1 *SE*.

The analysis was then repeated but with floor performers removed. The proportions of targets identified were calculated for replication and concealment combined: 27 older adults scored at or below chance. This was 31% of the 88 older adults (not taking into account the two older adults who were excluded earlier) and they were excluded from the following analysis. In order to exclude the same amount of young adults, the 31% of poorest young performers (as measured by proportion of targets identified in replication and concealment lineups combined) were also excluded from the following analysis (i.e., 18 out of 60). A similar exclusion technique has been used in previous research (e.g., Naveh-Benjamin et al., 2009). With the reduced data set, young adults still performed better than older adults, $F(1, 101) = 20.80$, $MSE = 0.09$, $p <$

.001. Performance was still better in replication lineups compared to concealment lineups, $F(1, 101) = 11.65$, $MSE = 0.12$, $p < .001$, and the interaction between age and lineup type remained, $F(1, 101) = 6.29$, $MSE = 0.12$, $p < .05$. Young and older adults' performance difference between replication and concealment was .29 and .04 respectively. Paired *t*-tests between replication and concealment lineup performance confirmed that young adults benefited from replication, $t(41) = 4.05$, $p < .001$, whereas older adults did not, $t < 1$. This analysis therefore demonstrates that it was not floor performance in older adults that was driving the interaction between age and lineup type. This further supports the earlier conclusion that older adults do not benefit from replication lineups as much as young adults, if at all.

There was also a possibility that the difference in benefit for replication over concealment between young and older adults was due to overall memory performance. The difference in the proportion of target endorsements between replication and concealment lineups was calculated and compared to two independent measures of memory performance. The first independent measure of memory performance was the proportion of none responses in all target-absent lineups and the second independent measure of performance was the proportion of none responses in all target-present lineups. Neither of the two measures correlated with the difference in performance between replication and concealment lineups, $r(148) = -.002$, $p = .98$, $r(148) = -.050$, $p = .54$, respectively. Therefore, the benefit of replication lineups over concealment lineups (as measured by correct target endorsements) is not determined by overall memory performance.

Modelling. A hybrid similarity model was constructed as outlined in Nosofsky and Zaki (2003). The similarity between a given lineup-face, study-face, pair was determined on the basis of four parameters *s*, *M* , *C* and *D*. The parameter *s* ($0 \leq s \leq 1$) represents the average *similarity* between pairs of non-identical faces ignoring any distinctive features. The parameter *M* ($0 \leq M \leq 1$) represents a reduction in similarity between a pair of faces when a distinctive feature is present in one face and *missing* in the other. The parameter *C* ($C > 1$) represents a boost in similarity between a pair of faces when they both share an identical *common* distinctive feature. The parameter *D* ($0 \leq D \leq 1$) represents a reduction in similarity between two faces that have a *different* distinctive feature. Table 5 shows how the parameters were combined to represent the overall similarity of a given lineup member to all faces in memory from the study period. For example, a target face in a replication lineup *S(TARGET)* has a distinctive feature that is not present (*missing*) in 26 of the originally studied faces, and that is *different* to other distinctive features in five of the originally studied faces. It is also identical to the same face at study (so $s = 1$) and therefore shares a *common* distinctive feature with that face.

Table 5

Average Similarity Between a Given Lineup Face and all of the Faces in Memory from the Study Set

Lineup Type	Similarity of Target, $S(TARGET)$	Similarity of a Foil, $S(FOIL)$
Replication	$26sM + 5sD + C$	$26sM + 5sD + sC$
Concealment	$26s + 5sM + M$	$26s + 5sM + sM$

For modelling, the similarity measures from Table 5 were used to calculate the probability of a participant making a *target*, *foil* or *none* response for target-present lineups and a *foil* or *none* response for target-absent lineups. The probability of making a given response was determined by five equations that compare the magnitude of target and foil similarities to faces in memory from the study period (i.e., the probability of endorsing a target is increased when the target face has a higher similarity to faces in memory and when foil faces have a lower similarity to faces in memory). A fifth parameter k ($k > 0$) was used to adjust the probability of making a *none* response (larger values of k increase the probability of making a *none* response). For target-present lineups the following three equations were used:

$$prob(TARGET) = \frac{S(TARGET)}{S(TARGET) + N_{Foils}\ S(FOIL) + k} \qquad (1)$$

$$prob(FOIL) = \frac{N_{Foils}\ S(FOIL)}{S(TARGET) + N_{Foils}\ S(FOIL) + k} \qquad (2)$$

$$prob(NONE) = \frac{k}{S(TARGET) + N_{Foils}\ S(FOIL) + k} \qquad (3)$$

For the target-absent lineups the following two equations were used:

$$prob(FOIL) = \frac{N_{Foils}\ S(FOIL)}{N_{Foils}\ S(FOIL) + k} \qquad (4)$$

$$prob(NONE) = \frac{k}{N_{Foils}\ S(FOIL) + k} \qquad (5)$$

N_{Foils} represents the number of foils in a lineup (N_{Foils} = 5 for target-present lineups and N_{Foils} = 6 for target-absent lineups).

The model was fit to all of the data (considering each participant's individual responses) using maximum likelihood to estimate the values for the five parameters *s*, *M*, *C*, *D* and *k* as described in Lamberts (2005). A restricted model was created by fixing the five parameters to be the same for young and older adults. This was then compared one at a time to five different general versions of the model. Each of the general versions had a different one of the five free parameters free to vary between young and older adults. A fully general model with all parameters free to vary between young and older adults was also fit to the data, as well as a general model where both *s* and *C* were free to vary between young and older adults. Table 6 shows the parameters for the restricted and general models.

Table 6

Restricted and General Model Parameters for Young (Y) and Older (O) Adults

Parameters free to vary by age group	Age group	Model parameters					log-likelihood	χ^2 comparison to restricted model
		s	*M*	*C*	*D*	*k*		
None (Restricted model)		0.01380	1.000	2.475	1.000	1.094	-1479	
s	*Y*	**0.00941**	1.000	2.529	1.000	1.084	-1462	$\chi^2(1) = 34.28$***
	O	**0.01780**						
M	*Y*	0.01265	**0.585**	1.999	1.000	0.9049	-1473	$\chi^2(1) = 11.57$***
	O		**1.000**					
C	*Y*	0.01317	0.994	**4.646**	1.000	1.043	-1466	$\chi^2(1) = 25.85$***
	O			**1.375**				
D	*Y*	0.01403	1.000	2.406	**0.000**	1.078	-1474	$\chi^2(1) = 8.80$**
	O				**1.000**			
k	*Y*	0.01380	1.000	2.477	1.000	**1.299**	-1475	$\chi^2(1) = 7.00$**
	O					**0.967**		
s & *C*	*Y*	**0.00979**	0.994	**4.159**	1.000	1.074	-1454	$\chi^2(2) = 49.42$***
	O	**0.01685**		**1.584**				
s, *M*, *C*, *D* & *k*	*Y*	**0.00738**	**0.742**	**2.389**	**0.949**	**0.665**	-1454	$\chi^2(5) = 48.70$***
	O	**0.02246**	**0.998**	**2.363**	**0.979**	**1.453**		

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

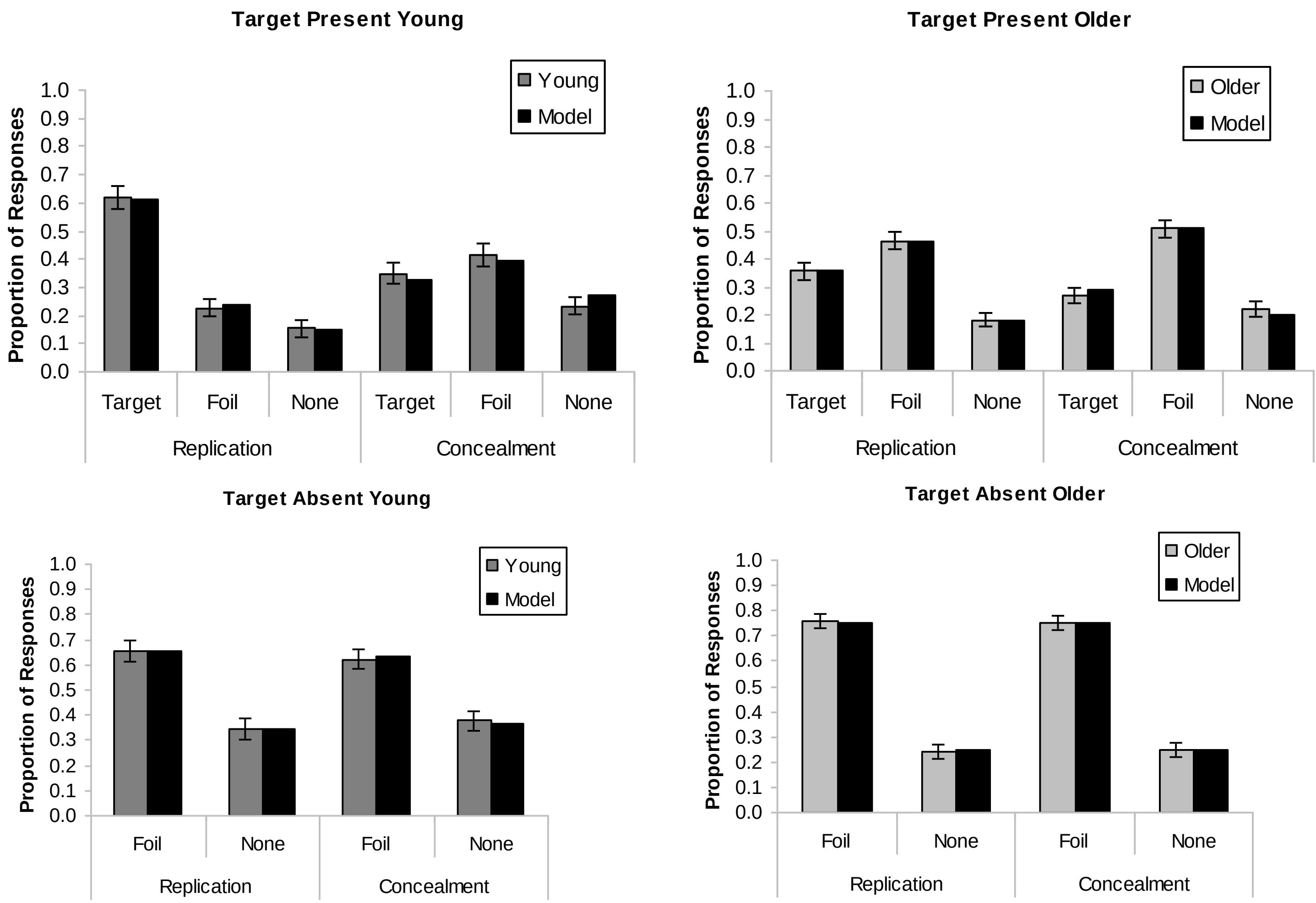

Figure 14. Mean proportion of responses in each response category (identifying a target, foil or none of the faces) in replication and concealment lineups for young (left) and older (right) adults and for target-present (top) and target-absent (bottom) trials. Error bars are ± 1 *SE*. Model data shows fits with parameters *s* and *C* free to vary between young and older adults.

The following formula from Lamberts (2005) was used as a likelihood-ratio test to establish if the general models were significantly better fits than the restricted model:

$$\chi^2 = -2\left[\ln L(restricted) - \ln L(general)\right] \quad (6)$$

Where lnL(restricted) and lnL(general) are the log-likelihood values of the restricted and general models, respectively. The degrees of freedom of the χ^2 statistic are the number of extra free parameters in the general model compared to the restricted model. The general model fit the data better than the restricted model when any single parameter was free to vary between young and older adults (see Table 6). As *s* and *C* individually improved the fit more than other parameters when allowed to vary between young and older adults, a general model with both *s* and *C* free to vary between young and older adults was tested against the restricted model and it was also a significantly better fit. Additionally, a fully general model with all parameters free to vary between young and older adults was not a better fit than the general model with just *s* and *C* free to vary between young and older adults (the fit was no better *at all*, not just statistically no better). The general model with *s* and *C* free to vary between young and older adults was also a significantly better fit than the general model with just *s* free to vary between young and older adults, $\chi^2(1) = 15.14$, $p < .001$, and the general model with just *C* free to vary between young and older adults, $\chi^2(1) = 23.57$, $p < .001$. This indicates that *s* and *C* are both important and are each making their own independent contribution to the fit of the model to the data. Figure 14 shows the model fits for when *s* and *C* were both free to vary between young and older adults against the observed data.

Discussion

The overall pattern in the data was consistent with the findings from Zarkadi et al. (2009) with replication resulting in superior target identification compared to concealment but without leading to a corresponding increase in foil identification in target-absent lineups. Older adults showed poorer memory performance than young adults and benefited from replication over concealment significantly less than young adults: Unlike young adults, older adults did not show any measurable difference in performance between the two lineup methods and this could not be explained by differences in overall memory performance between young and older adults. Even though older adults did not benefit from replication over concealment, replication did not hinder their memory performance. This means that the replication technique should still be recommended to police officers as a more suitable method compared with concealment of constructing lineups for suspects with distinctive features.

In the current study, for young adults, the overall level of performance was higher than in Zarkadi et al. (2009) for target-present lineups. This is easily explained by the fact that during the study period, faces were presented for 3 s each in the current study rather than 2 s each in Zarkadi et al.'s (2009) study. Also the delay between study and test was 90 s in the current study and 5 minutes in Zarkadi et al. (2009). For target-absent lineups, the level of performance was similar across the two studies, which is unusual considering the current study should have shown improved performance compared to Zarkadi et al. (2009). Older adults performed worse than young adults overall, which is consistent with research showing age-related decline in facial memory (e.g., Grady et al., 1995; Naveh-Benjamin, Guez, Kilb et al., 2004). Older adults' reduced facial memory may also have been due to that fact that the faces used in the study were all of young adults. There is evidence for an own-age bias in facial recognition, where participants are not as good at recognising faces of different ages to themselves (e.g., Perfect & Moon, 2005; Wright & Stroud, 2002). This could have provided a disadvantage to older adults.[15] In general, levels of performance were unlikely to have affected the difference between replication and concealment lineups because the difference in performance between the two lineup methods did not correlate with independent measures of memory performance.

An important applied aspect of this research must also be considered. Although replication appears to be superior to concealment as a method for creating fair lineups for suspects with distinctive features, it is not always the best method to use. Wells et al. (1998) argue that foils in replication style lineups should only possess features that the witness has reported to police officers. If, however, the suspect has a strongly distinctive feature that the witness has not reported, then they may still stand out in a lineup against non-distinctive foils. In this situation it may be more appropriate to use concealment. This would prevent the witness from identifying the suspect on the basis of a feature that they may have forgotten to report to police officers but that they still remember. Concealment could also prevent the witness from rejecting a face on the basis of it having a feature that they do not remember.

It is well established in the literature that distinctive stimuli are easier to memorise than non-distinctive stimuli (Schmidt, 1991) and the success of the hybrid similarity model with recognition of distinctive features makes it suitable for modelling facial identification (Knapp, Nosofsky, & Busey, 2006). It is a modified

[15] There was no evidence of an own-gender bias in the data, where participants show better memory for faces that are the same gender as themselves (e.g., Wright & Sladden, 2003). Male and female participants were equally successful at identifying the male faces and this did not interact with the difference in performance between replication and concealment lineups.

global familiarity model which is able to take into account distinctive features of memory stimuli (Nosofsky & Zaki, 2003). This allows the hybrid similarity model to predict greater recognition memory performance for distinctive memory stimuli compared to non-distinctive stimuli where the standard global familiarity model does not (Nosofsky & Zaki, 2003). In the current study, memory for non-distinctive faces was not tested so it may be interesting for future research to also include non-distinctive faces as targets to see how this affects the model's parameters.

The hybrid similarity model fit the overall pattern of data well for both young and older adults (see Figure 14). The model was able to predict superior performance for replication lineups compared to concealment lineups because in replication lineups the targets are more distinguishable from foils. The model fit the data best when parameters s and C were free to vary between young and older adults. The parameter s represents the average similarity between two non-identical faces; it is applied to all comparisons between pairs of faces regardless of similarities or differences in distinctive features. Older adults had a higher value of s than young adults, resulting in an increased overall familiarity of both targets and foils in the model. This corresponds to the higher levels of endorsement responses compared to *none* responses in older adults compared to young adults in the data for target-absent trials. The data are in line with studies that show increased levels of false memory in older adults compared to young adults (e.g., Castel & Craik, 2003; Norman & Schacter, 1997) and is particularly in line with studies that show increased levels of false identification of faces in older adults compared to young adults (e.g., Memon & Bartlett, 2002; Memon et al., 2004; Memon, Hope, Bartlett, & Bull, 2002). The higher level of s for older adults could also be related to the own-age effect outlined above, where faces of different ages to participants are less discriminable (i.e., more alike) from each other.

The parameter *C* represents a boost in similarity between a study item and a test item that share the same distinctive feature. The parameter *C* therefore only applies to replication lineups where test items have distinctive features. In concealment lineups, *M* represents the reduction in similarity between two items when one of the items has a distinctive feature that is missing in the other (e.g., the target face in a concealment lineup is missing a distinctive feature when compared to the same face at study). Older adults had a lower value of *C* than did young adults and it was much closer to the value of *M* (see Table 6). This corresponds to the fact that there was little difference between replication and concealment lineups in older adults.

There were three parameters that correspond to distinctive features in the model (*M*, *C* and *D*) but only parameter *C* was different to one. This means that *C* is the parameter responsible for the different effect of distinctive features in replication and concealment lineups: Parameter *C* represents a *boost* in similarity when a test and study face have matching distinctive features and only applies to replication lineups. Parameters *M* and *D* represent *reductions* in similarity between a pair of faces when distinctive features are present on one face and not on the other (*M*) or distinctive features are different on both faces (*D*). This indicates that the benefit of replication over concealment in young adults is driven by a boost in familiarity when a target is more congruent to an item in memory in replication lineups, rather than a decrease in familiarity when a distinctive feature is removed from a target in concealment lineups. It is therefore possible that if older adults form weaker associations between faces and distinctive features, then they cannot benefit from this boost and do not benefit from replication compared to concealment. This is what the model suggests as the value of C is much closer to one (no boost) in older adults.

Ultimately, the age-related associative deficit hypothesis predicts that older adults would form weaker links between faces and distinctive features. Forming an association between a face and a distinctive feature should have minimal impact on concealment lineups because that distinctive feature is not present as a cue. Therefore, age-related associative deficits are more likely to occur in replication lineups where associative memories are important (it can be seen in Figure 13 that age differences for replication are much larger than for concealment). Also the modelling process indicated that the main benefit of replication compared to concealment is due to boosts in familiarity of replication targets, not reductions in familiarity of concealment targets. Crucially, the current results indicate that age-related associative deficits not only impact on overall memory performance, but they can also influence the qualitative pattern of older adults' behavior.

Chapter 5: Age-Related Associative Deficits Are Absent with Nonwords

Chapter 2 discussed that age deficits in associative memory may arise from general deficits in item memory. In order to remember an association between two items, a person may need to have memory for those items individually. Therefore, if older adults have poorer item memory than young adults, this deficit is enhanced cumulatively for associative memory, which requires memory for multiple items (see Chapter 2 for more detail). However, it may be more appropriate to express item memory in terms of associative memory. Item memory itself requires associations; for example, in word memory a participant must associate visual patterns into letters and letters into words. Once the word is comprehended, it then must be associated to the experimental context in order to be correctly recalled or recognised later. Therefore, item memory itself may be affected by associative deficits and the increased inter-item associative deficits observed in older participants compared to young participants may simply be a magnification of the same effect.

The current chapter aims to explore this associative definition of item memory by manipulating preexisting knowledge related to items (i.e., the novelty of items). The majority of studies demonstrating age-related associative deficits employ an item test for familiar information and an associative test for novel information. The current study therefore aimed to explore age differences in item vs. associative memory by directly manipulating the novelty of individual items. This would clarify the distinction between item and associative memory by eliminating the difference in preexisting knowledge between item and associative memory.

Preexisting Knowledge

The increased age-related deficits observed for associations between items may be because inter-item associations are different to within item associations, namely, that they are completely new connections as opposed to preexisting concepts reactivated. Indeed, this is the view proposed by Naveh-Benjamin et al. (2003) who stated that age-related associative deficits are most apparent in the formation of completely new associations. For example, a typical associative memory measure is to present pairs of unrelated words and test for memory of the words themselves and also their pairings (e.g., Naveh-Benjamin, 2000, Exp. 2). The words (items) will have been seen by participants many times before and are therefore preexisting in memory; however, the pairings of unrelated words are novel and unique associations. If item memory itself is in fact purely associative in nature, it would seem that superior item memory compared to associative memory (as is found in both young and older participants, e.g., Old & Naveh-Benjamin, 2008a) may be due to reinforcement of item memory with preexisting concepts. Therefore, it is hypothesised that older adults benefit more from preexisting knowledge, resulting in reduced age deficits with item tests.

Preexisting knowledge has been shown to reduce age deficits in associative memory. In Naveh-Benjamin’s (2000) Experiment 4, young and older participants’ associative memory was tested for semantically related and unrelated pairs of words. For both cued and free recall, there were significant age-related associative deficits for unrelated pairs but not for related pairs. This indicates that older participants were able to dramatically improve their performance when preexisting semantic information could be used to support their associative memory.

The semantic relationship between pairs of words was further examined in the context of age-related associative deficits by Naveh-Benjamin et al.'s (2003) Experiment 2. Young and older participants were presented with pairs of words; half of the pairs consisted of two words that were semantically related to each other and half of the pairs consisted of two unrelated words. When word pairs were semantically unrelated, older participants had a poorer memory for associations between words than for the words themselves, whereas young participants showed equivalent memory for individual words and pair associations. However, when the word pairs were semantically related, both young and older participants showed equivalent memory for words and associations. Naveh-Benjamin et al. (2003) used this evidence in support of the argument that associative deficits in older participants are specific to new associations that are not supported by preexisting knowledge. A similar pattern of results was found by Patterson et al. (2009). Young and older participants were shown pairs of semantically related and unrelated words and age differences in associative memory were smaller for the related pairs.

A study by Castel (2005) found that older participants had particular deficits compared to young participants when memorising the association between objects and their prices when the prices were unusual. In Castel's (2005) Experiment 2, young and older participants were shown everyday groceries and corresponding prices which could be either normal market value, high or low. Participants were explicitly told to remember the prices of each object for a later memory test. Their memory was tested with cued recall by showing the objects and asking participants to recall the corresponding prices. The results showed no age differences for objects priced at market value, but older participants were significantly poorer than young participants at memorising over-priced and under-priced items. This experiment showed that the formation of unusual associations can be particularly difficult for older participants but when preexisting real world knowledge is available to support memory, age differences can be reduced.

Similarly, Castel (2007) examined how semantic relatedness influences arbitrary associations in young and older adults. Participants saw three-element phrases that consisted of a number, an object and a location. The numbers were always arbitrary but the objects and locations could be related (e.g., 86 hotels in the city) or unrelated (e.g., 58 nails in a bowl). The memory test consisted of cued recall where the location was given and participants were required to recall the number and object. Older adults were worse overall but age differences were largest for unrelated associations: Older adults showed only a small memory deficit for memorising related objects and locations but age deficits were larger for unrelated objects and locations and for numbers of objects.

This pattern of results was also found with one study of picture memory. Hess and Slaughter (1990) manipulated preexisting knowledge with images. The study tested the age differences related to associating objects to spatial locations. Young and older participants' memory for objects and their spatial positions was tested with organised and unorganised object positions: Illustrated objects (e.g., a sink) were placed within a scene (e.g., a kitchen). Organised scenes placed the objects in realistic positions (e.g., a shelf on the wall) while unorganised scenes were created by placing objects in unrealistic positions (e.g., kitchen shelf below kitchen sink, calendar on floor etc). Participants were shown the various scenes and their memory for both objects and their locations was later tested. Results showed that older participants were more reliant on preexisting knowledge when memorising object positions. For both young and older adults, organised scenes resulted in better object location memory than unorganised scenes. However the effect of organisation was more extreme for older adults. Overall, this study indicated that older adults rely more on preexisting knowledge to support memory than do young adults.

A similar experiment was conducted by Gutchess and Park (2009, Exp. 3) but the findings were different to those of Hess and Slaughter (1990). Groups of young and older participants looked at a series of images with a central picture (e.g., a cow) presented in front of a regular (e.g., farm) or irregular (e.g., laundry room) background. They were instructed to remember the central pictures and the background in front of which they were presented (i.e., associative memory). Following this, they completed a recognition memory test where some of the central pictures and backgrounds were mixed up and they had to respond yes/no as to whether each test image had been seen in that combination before. There was a main effect of relatedness at encoding with regular scenes being recalled better than irregular. There was also a main effect of age with young participants showing better recognition than older participants. However, there was no interaction between regularity and age and Gutchess and Park (2009) argued that older participants did not have a specific deficit for unusual associations for complex image memory.

Preexisting knowledge has also been shown to affect age deficits in source memory: Mather, Johnson and De Leonardis (1999) examined young and older participants' source memory whilst manipulating how stereotypical the relationship was between the source and the content. Statements were played to participants via video: A woman previously identified as for example, an athlete, could present a stereotypical statement (e.g., I enjoy competing in athletic events), a non-stereotypical statement (e.g., Writing is my passion in life) or a neutral (neither consistent nor inconsistent) statement. Participants had to later associate statements to speakers. When the statements were stereotypical or neutral there were no age differences in memory for source. However, when the statements were inconsistent with the source, older participants performed significantly worse than young participants. This experiment indicates that older participants were likely to use preexisting knowledge to remember associations, which was detrimental to associations inconsistent with that knowledge. An interesting point to note about these results is that older participants were no worse at memorising neutral statements. These neutral statements would not have been supported by preexisting knowledge and this indicates that in this case the use of preexisting concepts is not responsible for reducing age differences. Therefore, the use of preexisting knowledge was not necessarily supporting older participants' memory of stereotypical associations; rather it was detrimental to non-stereotypical associations.

Novel Stimuli

Several ageing studies have demonstrated associative deficits in older adults with unfamiliar faces, which are novel stimuli (Bastin & Van der Linden, 2005; Bastin & Van der Linden, 2006; James et al., 2008; Naveh-Benjamin, Guez, Kilb et al., 2004; Naveh-Benjamin et al., 2009; Rhodes et al., 2008). Only four of these studies had comparable measures of both item and associative memory and were therefore able to demonstrate larger associative memory deficits than item memory deficits in older adults: Naveh-Benjamin et al. (2004, 2009) and Bastin and Van der Linden (2005) did not associate the faces to novel stimuli so only half of the stimuli were novel; Bastin and Van der Linden (2006) used pairs of unrelated faces, but item memory age differences were probably restricted by ceiling effects (proportion correct for item recognition was 0.93 for both young and older adults), making the comparison between item and associative memory age deficits difficult to interpret.

A more important point is that none of these studies directly tested the effect of stimulus novelty on age-related associative deficits – that is, previous work did not compare both item and associative memory age differences between novel and familiar stimuli.

Of most relevance to the current study is Naveh-Benjamin's (2000) Experiment 1, where word-nonword pairs produced age-related associative deficits: Age deficits in an associative memory test for word-nonword associations were significantly larger than age deficits in an item test. Interestingly, the age-related associative deficit was smaller when nonwords were used as an item test compared to when words were used (age differences of .03, .23 and .36 for words, nonwords, and word-nonword associations, respectively), suggesting that the use of nonwords may have an influence on the size of age-related associative deficits.

Experiment 4

In the current study, nonwords were used to remove the support of preexisting knowledge for item memory. This would mean that when item memory and associative memory were compared, the memories formed would be equally novel across tests. Naveh-Benjamin et al.'s (2003) Experiment 2 reduced the novelty of associative memories by using semantically related words when comparing item and associative memory. The current study complements that design by increasing the novelty of the item memory measure in order to better compare it with the associative memory measure. In addition, the standard age-related associative deficits were reproduced with words for comparison to the nonwords data. Thus the study was designed to directly investigate the effect of stimulus novelty on age-related associative deficits.

Method

Both the words and nonwords conditions used the same general procedure as that of Naveh-Benjamin (2000). Pairs of stimuli were sequentially presented on a computer screen and were followed by separate item and associative recognition memory tests. Participants were explicitly instructed to remember the stimuli and the associations between them for a later recognition test and they completed a short practice session before the main procedure.[16] At encoding, pairs were presented in lower case in black font on a white background to the left and right of the centre of a computer screen, with a clear separation between them. Stimulus order was randomised at both encoding and test. At test, stimuli were randomly assigned to either the item or associative tests. In addition, all word pairs were semantically unrelated.

Nonwords condition. Memory set sizes and stimulus presentations were refined in three pilot studies that were conducted in order to avoid floor and ceiling effects in memory performance.

Pilot study 1. In the first pilot study, the practice consisted of six pairs of nonwords; each pair was presented for 6 s and a 1-minute distracter delay period followed the sequence before the memory test. During the delay period, participants were instructed that they would need to count out loud backwards in threes from 200. Following the delay, there were four item tests and four associative tests, each test having two old and two new stimuli for response. For the item test, a single nonword was presented on the screen and participants were asked to press the left mouse button if they had seen it before or the right mouse button if it looked

[16]During the practice, many participants in the nonwords condition thought that the test was going to be too difficult as they naturally expected a free recall style test. After completing the recognition practice test, participants were more confident. Therefore, the practice served not only to familiarise participants with the task but also to prevent them from giving up when trying to memorise the larger main set of nonwords.

completely new. Only once the button was pressed did the next word appear so the test was self paced. The associative test was similar: Two nonwords were presented on the screen and participants were informed that both of the nonwords had definitely been seen before. Participants were then told to press the left mouse button if they thought that the two nonwords were originally together in a pair, or to press the right button if they thought one non-word was from one pair and the other was from a different pair.

The main memory set consisted of 26 pairs of words with the first and last pairs used as buffers. The remaining 24 pairs were to be tested. As in the practice, each pair was presented for 6 s with a 1-minute distracter delay period at the end of the sequence. The memory test consisted of 32 item tests (16 old and 16 new items) and 16 associative tests (8 old and 8 recombined pairs). The counting task and the tests were the same as for the practice. For both the main and the practice tests, the test type was counterbalanced so that half of the participants in each age group received the item test before the associative test and vice versa for the other half.

Ten participants (6 female) of all ages (M = 43.9 years, SD = 23.67, range 20-82) completed the first pilot study.

For the item test, a Wilcoxon Signed-ranks test indicated that there were significantly more hits (M = 11.3, SD = 2.6) than false alarms (M = 4.4, SD = 2.3), Z = 2.81, $p < .01$. Therefore participants demonstrated memory for the items as they responded more successfully than if they were responding randomly. No participants scored 100% on the item test so ceiling effects were not present. For the associative test, a Wilcoxon Signed-ranks test indicated that there was no significant difference between the number of hits (M = 4.6, SD = 1.4) and false alarms (M = 4.1, SD = 1.2),

$Z = 1.07$, $p = .29$. Therefore, participants demonstrated no memory of the associations as their responses were no better than chance.

The results showed that there were floor effects for the associative test as it proved to be too difficult. The second pilot study aimed to rectify that with reduced task difficulty.

Pilot study 2. In order to reduce the difficulty in the second pilot study, the memory sets were presented twice and the main study list was shortened slightly.

The practice still consisted of six pairs of nonwords. This time the pairs were each presented for 5 s sequentially. Following the presentation of the memory set, a screen appeared for 10 s indicating that 'The repeated showing of nonwords will begin shortly'. Following this, the pairs were presented again in exactly the same order and rate. A 1-minute distracter delay period before testing was used (with backwards counting from 200) and there were four item tests and four associative tests as in the first pilot study.

The main memory set consisted of 23 different pairs of nonwords. The first pair was a buffer and this was followed by 21 pairs that would be later tested. After the 21 nonword pairs had been shown, a screen indicated that the list would be repeated and the 21 pairs were shown again. After this, a final buffer pair of nonwords was shown and then a 1-minute distracter delay period occurred before testing. As in the practice, each pair was presented for 5 s. The memory test consisted of 28 item tests (14 old and 14 new items) and 14 associative tests (seven old and seven recombined pairs). For both the main and the practice tests, the test type was counterbalanced so that half of the participants received the item test before the associative test and vice versa for the other half.

Six participants (4 female) of all ages (*M* = 45.5 years, *SD* = 29.30, range 16-77) completed the second pilot study.

For the item test, a Wilcoxon Signed-ranks test indicated that there were significantly more hits (*M* = 9.0, *SD* = 2.7) than false alarms (*M* = 1.8, *SD* = 1.9), $Z = 2.23$, $p < .05$. Therefore, participants demonstrated memory for the items as they responded more successfully than if they were responding randomly. As in the first pilot study, no participants scored 100% so ceiling effects were not present. For the associative test, a Wilcoxon Signed-ranks test indicated that there was no significant difference between the number of hits (*M* = 4.3, *SD* = 1.2) and false alarms (*M* = 3.5, *SD* = 1.6), $Z = 0.96$, $p = .34$. Therefore participants demonstrated no memory of the associations as their responses were no better than chance.

As the associative test performance was still low, the test difficulty was reduced in a final pilot study.

Pilot study 3. All aspects of the third pilot study were identical to the second except that the practice and main memory sets were repeated three times instead of two. There were two buffer pairs in the initial and final positions which were only shown once; the remaining 21 pairs of words were repeated three times (i.e., a buffer pair then 21 pairs repeated three times then a final buffer pair).

Six participants (3 female) of all ages (*M* = 58.2 years, range 24-85) completed the third pilot study.

For the item test, a Wilcoxon Signed-ranks test indicated that there were significantly more hits (*M* = 11.0, *SD* = 1.3) than false alarms (*M* = 2.0, *SD* = 0.9), $Z = 2.21$, $p < .05$. Therefore participants demonstrated memory for the items as they responded more successfully than if they were responding randomly. As before, no participants scored 100% so ceiling effects were not present. For the associative test, a Wilcoxon Signed-ranks test indicated that there was a marginally significant difference between the number of hits (M = 5.2, SD = 1.2) and false alarms (M = 3.3, SD = 1.2), Z = 1.84, p = .07.

The results indicated that floor effects would be unlikely in the associative test if the participant numbers were increased. Therefore the majority of the main experiment was conducted in exactly the same way as the third pilot study.

Main nonwords conditions procedure. The procedure was the same as that described for the second and third pilot studies. In brief, the main study involved the presentation of 21 pairs of nonwords. This was then followed by a distracter delay period (with backwards counting) of 1 minute. After the memory set presentation, participants completed item and associative recognition tests involving yes/no button presses for old/new stimuli.

There were two young groups of participants in the nonwords condition and one older group. One young group and the older group completed the nonwords memory test exactly as described for the third pilot study with three repetitions of the memory set. In order to produce data from young participants with a similar level of performance to the older group, a second group of young participants completed the nonwords memory test exactly as described for pilot study two. This meant that the second young group only saw the memory set two times and therefore the task was more difficult for them.

In the nonwords conditions, reaction times were collected during the recognition tests. Prior to testing, participants were not informed that their reaction times would be measured or that they needed to respond as quickly as possible. Therefore, the emphasis during testing was to make the correct choice and not to focus on speed.

Participants. A total of eighty participants took part in the study with 50 young and 30 older adults.

Thirty young adults (18 female), aged 16-31 years (*M* = 23.9, *SD* = 4.0), and 30 healthy older adults (22 female), aged 57-88 years (*M* = 74.3, *SD* = 8.4), took part in the experiment with three memory set repetitions. This includes three older participants who took part in the third pilot study. The remaining three participants from the third pilot study were excluded because they had previously taken part in an earlier pilot and their results may have been affected in an unpredictable manner. Young participants were recruited from the local community. Eight of the young participants received a financial incentive of £5 for taking part in the study; the remaining 22 young participants were offered no financial incentives. Older participants were recruited from the University of Warwick Age and Memory Study volunteer panel that was populated by local advertisements; they were offered no financial incentives for participation.

The remaining 20 young participants (16 female), aged 16-25 years (*M* = 19.9, *SD* = 2.2), completed the experiment with two memory set repetitions. This includes three participants who took part in the second pilot study. Young participants were recruited from the University of Warwick psychology department. Nine participants received £2 for taking part in the study and the remaining 11 received no financial incentives.

To assess cognitive functioning, participants completed the Digit Symbol Substitution task (Wechsler, 1981) as a measure of perceptual-motor or processing speed. They also completed the multiple choice part of the Mill Hill vocabulary test (Raven et al., 1988) as a measure of crystallised intelligence (see Table 7 for means). The results were consistent with the literature (e.g., Horn & Cattell, 1967; Salthouse, 1991). Young participants (all 50 together) were significantly faster at the Digit Symbol Substitution task than older participants, $t(78) = 10.45$, $p < .001$. For the vocabulary test, young participants scored significantly lower than older participants, $t(78) = 7.21$, $p < .001$. Further comparisons were made between the two young groups; the two-repetition young group were faster at the Digit Symbol Substitution task than the three-repetition young group, $t(48) = 2.16$, $p < .05$. This could be due to the different recruitment methods, but there was no performance difference for the Mill Hill vocabulary test between the two and three repetition young groups ($t < 1$).

Table 7

Mill Hill Vocabulary Test and Digit Symbol Substitution Task (DSST) Scores for the Participants from the Nonwords Conditions

Group	Mill Hill		DSST	
	M	*SD*	*M*	*SD*
Young (2 repetitions)	17.00	3.23	74.80	8.78
Young (3 repetitions)	17.30	3.16	67.63	12.96
Older	22.67	3.51	42.57	11.00

Materials. Ninety nonwords (see Appendix 3) were selected from the English lexicon project (Balota et al., 2007). The nonwords were chosen for specific characteristics: None of them had orthographic neighbours. They were all between six and nine letters in length (*M* = 8.21, *SD* = 0.87). They all had a high probability of being correctly identified as a nonword in a lexical decision task (*M* = 0.96, *SD* = 0.03, range = 0.90 -1.00). Also, to ensure the nonwords were not too different to normal words, they were selected to have a long reaction time when being judged as nonwords (*M* = 883 ms, *SD* = 51.3).

Nonwords were paired together and displayed on a laptop computer screen with a clear separation between the two items. The nonwords were presented centrally in black 32 point font with a white background. Participants viewed the nonwords at a distance of approximately 50 cm and the height of a word corresponded to approximately 1.15^{o} viewing angle. During a practise phase, participants viewed six pairs of nonwords and during the main phase, participants viewed a further 23 pairs of nonwords. From the main phase, the first and the last pairs were used as buffers so memory for these nonwords was not tested.

Recognition tests of memory were used. For tests of item memory, 28 nonwords were presented comprising 14 nonwords from the study phase and 14 new nonwords that were not previously presented. During the test, participants saw the nonwords one at a time at the centre of the screen and made a yes/no button press corresponding to an old or new nonword, respectively. For tests of associative memory, 14 pairs of nonwords were displayed sequentially. Seven pairs were from the study phase and seven were recombined pairs that consisted of nonwords presented during the study phase but not previously presented together. As with the item test, participants made a yes/no button press after each pair presentation to indicate old or new pairs of nonwords, respectively. A scaled down version was used for the practice phase; the item test was four trials with two old nonwords and two new nonwords. The associative test was four trials with two old pairs and two recombined pairs.

Randomisation of the stimuli was conducted separately for each participant. Fourteen nonwords were taken from the 90 to be used in the practice test. Any 12 of these items could appear in the six-pair practice study list. Of the 12 studied nonwords, any could appear as old nonwords or recombined nonwords in the practice item test. Also any of the six pairs could appear in the old pair part of the associative test. For the recombined pair associative test, any of the 12 previously studied nonwords could be presented together under the constraint that nonwords originally presented together could not appear together. Also, nonwords originally presented on the left remained on the left and nonwords originally presented on the right remained on the right. The remaining two nonwords of the 14 were presented as the new nonwords in the item test. Finally, there was an additional constraint such that no nonword could be presented twice during the recognition tests.

The remaining 76 nonwords from the 90 were used in the main test in much the same way. For the study phase, from the 76 nonwords, any four could be presented as the first or the last pair and would be labelled as buffers not to be tested in the recognition tests. From the remaining 72 nonwords, any 42 could appear in the 21 pairs of nonwords that would later be tested. For the item recognition test, any of the 42 previously presented items could appear as old or recombined nonwords. For the associative memory test, any of the 21 pairs could reappear as old pairs of associated nonwords. Also any two of the 42 studied nonwords that were not originally presented together could appear as recombined pairs in the associative test. As in the practice test, this was under the constraint that nonwords originally presented on the left were presented on the left during the recognition test and nonwords originally presented on the right were presented on the right in the recognition test. This left 30 previously not presented nonwords, any of which could appear as new nonwords in the item recognition test. Finally, as in the practice phase, there was a constraint that no nonword could be presented twice during the main recognition tests.

Words condition. In addition to the nonwords conditions, data were also acquired for words.[17] These data allowed assessment of how word and nonword memory interacts across age groups and memory type (item/associative). Similar to the nonwords condition, a sequence of word pairs was shown to groups of young and older participants and memory was then tested via item and associative recognition tests.

Two major differences between the words and the nonwords conditions (besides the stimuli) were as follows: The nonwords condition showed the study list three times to avoid floor effects whereas the words condition showed the study list just once (the task is easier with words so one showing was sufficient to avoid floor effects). The other major difference was due to the hypothesis being tested in the words condition; the original study aimed to identify the effect of time of day on age-related associative deficits. This meant that the words study was conducted twice for each participant, once at their optimal time of day (morning or evening) and once at their non-optimal time of day.[18] The order of testing was counterbalanced so that an equal number of young and older participants completed their first test in the morning as opposed to the evening and vice versa. Crucially, there was no significant overall practice effect from performing the task twice and no influence of practice on age-related associative deficits. Therefore, in order to compare the data with the nonwords condition, the results from the two periods of time were averaged for each participant to provide single measures of both item and associative memory.

[17] Data for the words condition were gathered by Laura Steel and Katherine Tyler and were originally part of a study that aimed to test age-related associative deficits at different times of day.
[18] The optimum time was determined by Horne and Ostberg's (1976) Morningness-Eveningness Questionnaire.

Participants. Twenty-four young adults (8 female), aged 18-25 years (*M* = 21.0, *SD* = 1.2), and 24 healthy older adults (15 female), aged 65-85 years (*M* = 75.3, *SD* = 6.3), took part in the experiment.

Materials. The words used in the experiment were provided by M. Naveh-Benjamin (personal communication to E. A. Maylor, November 3, 2008). The lexical characteristics of the words were not analysed in the original study so they were analysed independently here: The English lexicon project database (Balota et al., 2007) was used to assess certain characteristics of the words. Six different study lists were produced, each with corresponding item and associative tests. In total, 476 different words were used in the experiment. The words varied from 3-11 letters in length (*M* = 6.30, *SD* = 1.11). They occurred with a mean frequency of 8.75 (*SD* = 1.74, range = 3.73-12.99), using log HAL frequency (Lund & Burgess, 1996). Also the words had an average of 1.34 orthographic neighbours (*SD* = 1.91, range = 0-13). Word pairs were presented in black in the centre of a computer screen with the two words separated by a hyphen. Words were presented in a font size of 89 pt with a height corresponding to approximately 2° viewing angle at a distance of 60 cm.

Procedure. Before completion of the main words condition, each participant completed a practice test; this consisted of three word pairs presented sequentially, followed by a two-trial item test and a two-trial associative test. Before the memory set presentation, participants were explicitly informed about the nature of the task and were aware that memory would be tested later. Similarly to the nonwords condition, the item test showed a single word on the screen; one of the item test trials showed a word from the memory set and the other showed a previously unseen word. Likewise, the associative memory test was similar to the nonwords memory condition. For each trial, a pair of words was presented on the screen; one trial showed an intact pair from the memory set and the other showed a recombined pair where one word was from one pair and the other was from another. For the item test participants had to respond verbally yes or no to old or new words, respectively, and for the associative test participants responded yes to previously seen pairs and no to recombined pairs; the experimenter noted their responses on a mark sheet.

In the main experimental session, each participant viewed 34 word pairs sequentially. The first and last two pairs were buffers and memory was subsequently tested for the remaining 30 pairs. Each pair remained on the screen for 4 s – as in the practice, participants were explicitly instructed to study the pairs for a later memory test. After the memory set presentation and before the recognition tests commenced, a distracter delay period was completed in order to minimise recency effects. Participants were required to count backwards in threes from 300 for 60 s. After this, the recognition tests commenced; the format was the same as for the practice test described above. The item test had 40 trials, with 20 old words from the memory set and 20 previously unseen words. The associative test had 20 trials, with 10 intact word pairs from the memory set and 10 recombined pairs with two words from different pairs of the memory set. Participants were given as much time as they wanted to respond to each trial of the recognition tests. The experiment was counterbalanced so that half of the young and half of the older participants received the item test before the recognition test and vice versa.

Results

In order to run statistical analysis, the data were used to calculate the proportion of hits minus the proportion of false alarms people made. A hit was a correct positive response to a previously seen stimulus and a false alarm was an incorrect positive response to a previously unseen stimulus. This was done separately for both age groups (young and older), both test types (item and associative) and both conditions (words and nonwords). This created a uniform scale in which to compare each category of data, where chance performance gave a score of zero and perfect performance gave a score of one.

In addition to hits minus false alarms, the hit rates and false alarm rates were used to calculate d', which is a different way to assess performance based on signal detection theory. With d', the separation between the signal (old stimuli) and noise (new stimuli) of a response is represented in units of the standard deviation of the noise distribution (see Snodgrass & Corwin, 1988, for more detail). However, d' is particularly sensitive to extreme values where either the hit rate or the false alarm rate is at zero or one; in these situations the z scores used to calculate d' tend towards $-\infty$ or $+\infty$, respectively, and d' tends towards $\pm\infty$. Where this occurred in the data, a rate of zero was replaced with $0.5/n$ and a rate of one was replaced with $(n-0.5)/n$ where n is the number of old/new trials. This correction is attributed to Macmillan and Kaplan (1985) and is the most commonly used solution for extreme values (Stanislaw & Todorov, 1999). A further statistic, β, was also calculated to assess response bias. For analytical purposes ln(β) was used; ln(β) provides a negative value for bias towards yes/seen before responses and a positive value for bias towards no/not seen before. Therefore a lack of any bias produces a zero value of ln(β) (see Snodgrass & Corwin, 1988, for more detail).

To assess any effects of test order (whether the item test was before the associative test or vice versa), the results were entered into a 4-way repeated measures ANOVA. A 2 (Age: young,[19] older) x 2 (Memory test: item, associative) x 2 (Test order: item then associative/associative then item) x 2 (Condition: words,

[19] For the nonwords participants, only the three-repetition young group was included in this analysis.

nonwords) repeated measures ANOVA was conducted on the hits minus false alarms data. There was no main effect of test order, $F < 1$; also none of the interactions with test order was significant. Using the d' data, the main effect of test order was also non-significant, $F < 1$, and none of the interactions was significant. This indicates that the test order did not affect the results. The following analyses were therefore conducted without the test order factor.

Nonword memory. A 2 (Age: young 3 repetitions, older) x 2 (Memory test: item, associative) repeated measures ANOVA was conducted on the recognition data (hits minus false alarms performance) for the nonwords condition (see Figure 15 and Table 8 for means, and upper panel of Table 9 for ANOVA). There was a main effect of age, with young participants showing higher performance than older participants on average (*M* (*SD*) = 0.61 (0.21) and 0.38 (0.25), respectively). There was also a main effect of test type, with a higher performance in the item test than the associative test on average (*M* (*SD*) = 0.68 (0.17) and 0.31 (0.30), respectively). Finally, the interaction was non-significant. This shows that for nonwords there is no differential performance between item and associative memory across age, that is, age-related associative deficits were not present. The pattern of results was identical using the d' data (see Tables 8 and 9).

The same 2 x 2 analysis was conducted with the young participants who only saw two repetitions of the memory set (see Figure 15 and Table 8 for means, and upper panel of Table 10 for ANOVA). With these data there was no main effect of age. The other statistics showed the same pattern as with the young participants who saw the memory set three times. An important point to note is that there remained no test type by age interaction. This demonstrates that the lack of age-related associative deficits is not due to performance levels as there was no main effect of age in this comparison. In fact, the trend is in the opposite direction with larger age differences for the item test.

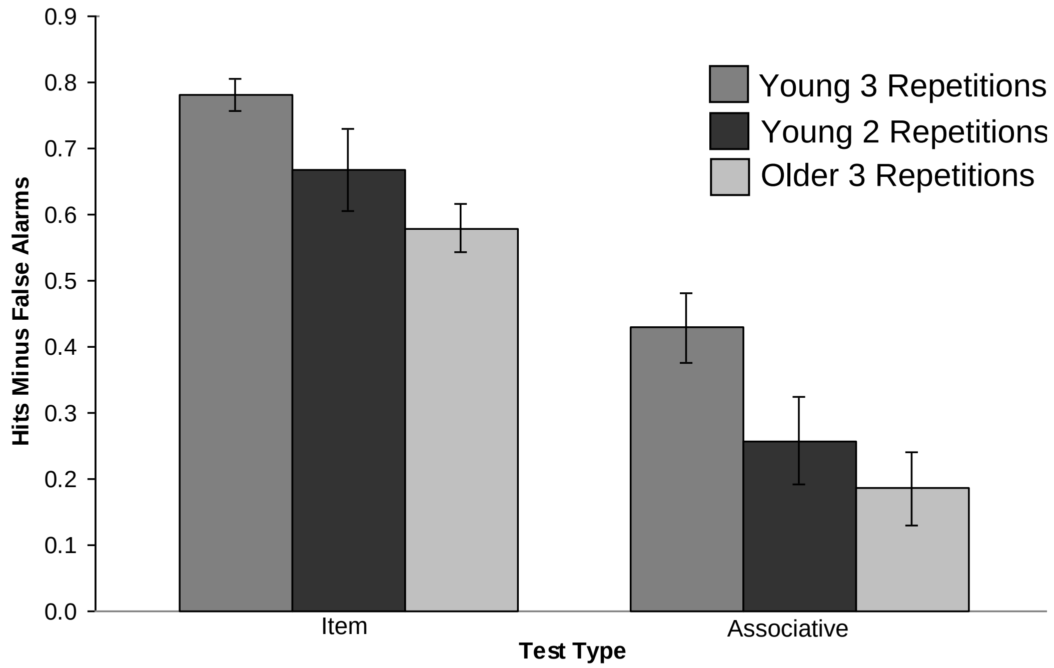

Figure 15. Hits minus false alarm performance for nonwords item and associative memory tests. Data are shown separately for young participants who saw 3 or 2 repetitions of the memory set and older participants who saw 3 repetitions of the memory set. Error bars are ± *1* SE.

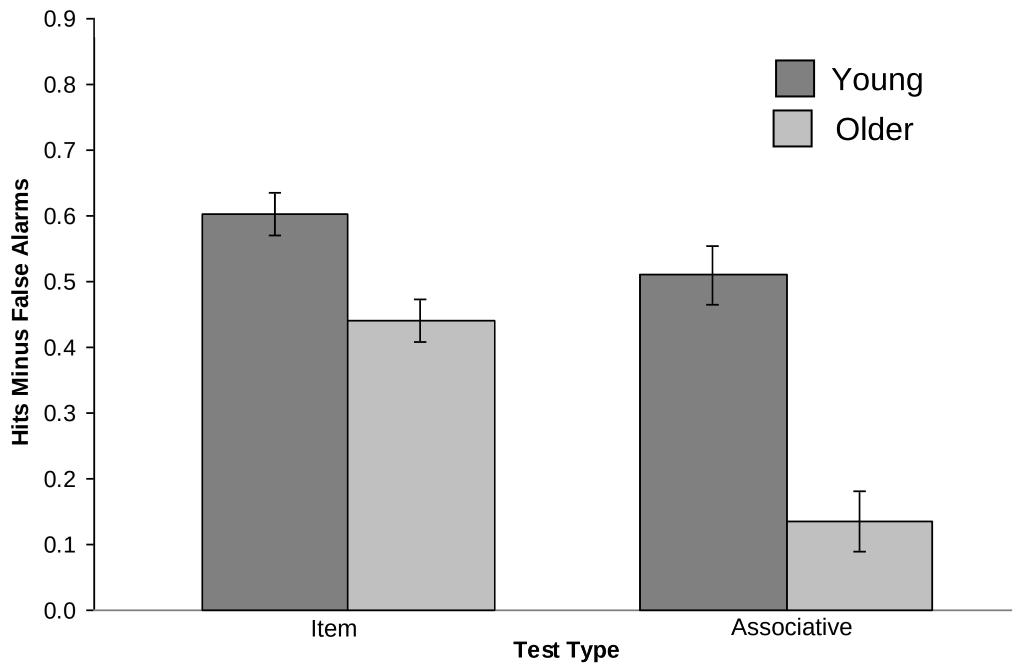

Figure 16. Hits minus false alarm performance for words item and associative memory tests. Error bars are ± 1 *SE*.

Table 8

Means and Standard Deviations for Item and Associative Recognition Memory Performance

Condition and Group	Item										Associative									
	H		FA		H-FA		d'		ln(β)		H		FA		H-FA		d'		ln(β)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Nonwords																				
Y*	0.89	0.08	0.13	0.10	0.78	0.13	2.58	0.56	-0.10	0.86	0.72	0.15	0.31	0.19	0.43	0.29	1.23	0.86	-0.01	0.49
Y**	0.83	0.15	0.18	0.17	0.67	0.28	2.21	1.05	0.03	0.71	0.66	0.16	0.42	0.22	0.26	0.30	0.73	0.82	-0.02	0.44
O	0.75	0.14	0.18	0.14	0.58	0.20	1.81	0.73	0.32	0.71	0.60	0.23	0.41	0.20	0.19	0.30	0.53	0.87	-0.06	0.36
Nonwords reduced data[a]																				
Y*	0.90	0.07	0.10	0.07	0.82	0.10	2.72	0.48	-0.01	0.79	0.79	0.14	0.21	0.14	0.61	0.20	1.77	0.63	-0.01	0.62
Y**	0.85	0.14	0.15	0.14	0.71	0.24	2.37	0.92	0.01	0.75	0.71	0.12	0.30	0.17	0.43	0.23	1.20	0.67	0.01	0.54
O	0.77	0.12	0.18	0.14	0.60	0.19	1.86	0.73	0.33	0.75	0.67	0.19	0.31	0.16	0.37	0.19	1.07	0.56	0.00	0.38
Words																				
Y	0.70	0.15	0.10	0.09	0.60	0.17	1.98	0.65	0.82	0.64	0.71	0.18	0.20	0.14	0.51	0.26	1.57	0.91	0.20	0.52
O	0.61	0.13	0.17	0.11	0.44	0.14	1.34	0.55	0.56	0.61	0.60	0.17	0.46	0.21	0.14	0.17	0.37	0.48	0.04	0.22
Words reduced data[b]																				
Y	0.73	0.13	0.10	0.10	0.63	0.17	2.09	0.67	0.82	0.71	0.76	0.14	0.16	0.10	0.60	0.21	1.86	0.77	0.26	0.57
O	0.60	0.14	0.17	0.11	0.44	0.15	1.35	0.59	0.60	0.65	0.61	0.17	0.42	0.19	0.19	0.15	0.52	0.40	0.02	0.20

[a]Data with worst 40% of associative performers removed. [b]Data with the worst 21% of associative performers removed. H = Hit rate; FA = False alarm rate; H-FA = Hits minus false alarm rate. *Young 3 repetitions of the memory set nonwords data. **Young 2 repetitions of the memory set nonwords data.

Table 9

ANOVA Results Based on Hits Minus False Alarms (H-FA) and d' Data for Nonwords

Effect type		H-FA		d'	
	df	*F*	*MSE*	*F*	*MSE*
Full data					
Age	1	20.36***	0.07	20.93***	0.78
Memory test	1	100.82***	0.04	131.92***	0.39
Age x Memory test	1	0.30	0.04	0.10	0.39
Error	58				
Reduced Data[a]					
Age	1	19.88***	0.05	19.16***	0.58
Memory test	1	56.03***	0.02	83.62***	0.16
Age x Memory test	1	0.12	0.02	0.64	0.16
Error	34				

Note. Young data includes only participants who saw 3 repetitions of the memory set.
[a] Data with the worst 40% of associative performers removed (12 from each age group).
*** $p < .001$.

Table 10

ANOVA Results Based on Hits Minus False Alarms (H-FA) and d' Data for Nonwords

Effect type		H-FA		d'	
	df	*F*	*MSE*	*F*	*MSE*
Full data					
Age	1	1.71	0.09	2.34	0.95
Memory test	1	72.23***	0.05	85.64***	0.53
Age x Memory test	1	0.04	0.05	0.48	0.53
Error	48				
Reduced Data[a]					
Age	1	1.62	0.07	1.85	0.77
Memory test	1	41.67***	0.02	55.92***	0.25
Age x Memory test	1	0.56	0.02	2.02	0.25
Error	28				

Note. Young data includes only participants who saw 2 repetitions of the memory set.
[a] Data with the worst 40% of associative performers removed (8 from young and 12 from older age group).
*** $p < .001$.

Some participants scored at chance level during the associative tests. This would mean that the difference between item and associative memory would be reduced as associative performance hit a floor. As the nonwords condition showed no interaction between age and test type it was important to address whether or not this was due to floor effects. In order to remove any possible influence from floor effects, for the associative test, participants scoring at or below chance (based on hits-false alarms) were removed from the data. There were 12 older participants (40%) scoring at or below chance for the associative memory test. For the two and three repetition young groups, there were four and two participants, respectively, scoring at or below chance. In order to remove the same proportion of participants from each age group, the worst 40% of participants (from the associative test) were removed from both the young groups and the older group. A similar process to this was conducted by Naveh-Benjamin et al. (2009); in their experiment (involving both item and associative memory), participants scoring poorly on associative tests were removed and the results were unaffected. The above analysis was redone with the reduced data set. None of the analyses provided any qualitatively different results, indicating that the lack of age-related associative deficits was not due to floor effects. For a full comparison of results, see Table 8 for means, and lower panels of Tables 9 and 10 for ANOVAs.

In order to further test the dependence of associative memory on item memory, age differences in associative memory were assessed using ANCOVA with item memory as a covariate. The ANCOVA was conducted with the hits minus false alarms data for the three repetition young group with worst performers included.[20] There was no significant effect of age on associative memory when item memory was used as a covariate, $F(1, 57) = 2.35$, $MSE = 0.08$, *ns*, and the covariate, item memory, was significantly related to associative memory, $F(1, 57) = 6.55$, $MSE = 0.08$, $p < .05$. Levene's test for equality of error variances was non-significant, $F < 1$, indicating that the homogeneity of variance assumption had not been violated. This shows that for nonwords, age differences in associative memory can be accounted for by age differences in item memory. The same pattern of results was found using the d' data: There was no significant effect of age on associative memory when item memory was used as a covariate, $F(1, 57) = 1.99$, $MSE = 0.67$, *ns*, and the covariate,

[20] When comparing the two repetition young group to the older group, age differences were not present at all so this ANCOVA was not conducted with those data.

item memory, was significantly related to associative memory, $F(1, 57) = 7.67$, $MSE = 0.67$, $p < .01$.

In order to assess response bias, values of β were calculated from the data. A 2 (Age: young 3 repetitions, older) x 2 (Memory test: item, associative) repeated measures ANOVA was conducted on the ln(β) recognition data (see Figure 17 and Table 8). There was no main effect of test type, $F(1, 58) = 1.89$, *ns*, nor of age, $F(1, 58) = 2.21$, *ns*. However there was a significant interaction between test type and age, $F(1, 58) = 5.12$, $MSE = 0.33$, $p < .05$. This is because young participants' responses were relatively unbiased for both item and associative test types whereas older participants showed a bias towards no/not seen before responses for the item test but were relatively unbiased for the associative test. With the worst 12 associative test performers removed from each age group to avoid floor effects, the main effects of test type remained non-significant $F(1, 34) = 1.78$, *ns*, as did the main effect of age, $F < 1$, and the interaction became non-significant, $F(1, 34) = 1.76$, *ns*.

The β analysis was repeated with the two-repetition young group (see Figure 17 and Table 8). This time there was a main effect of test type, $F(1, 48) = 4.21$, $MSE = 0.27$, $p < .05$, but not of age, $F < 1$, and there was no interaction between test type and age, $F(1, 48) = 2.49$, $MSE = 0.27$, $p = 0.12$. With the worst 40% of associative memory performers removed (8 young and 12 older participants), there was no main effect of test type, $F(1, 28) = 1.34$, $MSE = 0.29$, *ns*, there was no main effect of age, $F < 1$, and no interaction between age and test type $F(1, 28) = 1.38$, $MSE = 0.29$, *ns*.

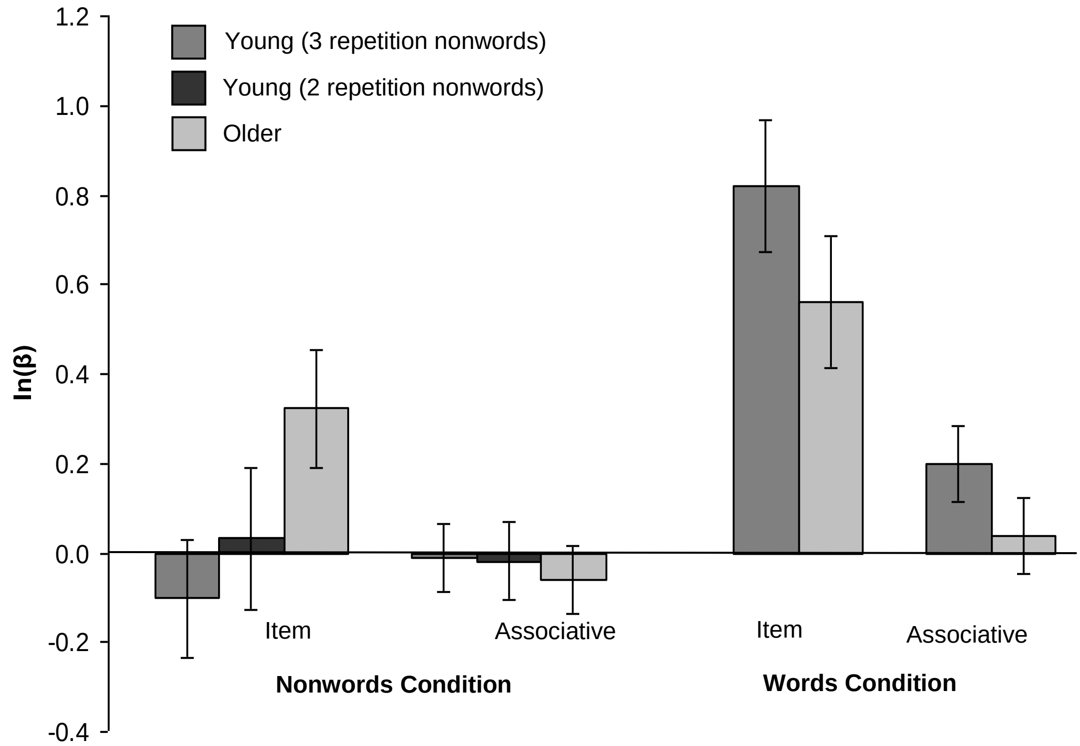

Figure 17. Response bias (ln(β)) for both nonwords and words conditions, item and associative tests and young and older participants. Error bars are ± 1 *SE.*

Word memory. A 2 (Age: young, older) x 2 (Memory test: item, associative) repeated measures ANOVA was conducted on the recognition data (hits minus false alarms performance) for the words condition (see Figure 16 and Table 8 for means, and upper panel of Table 11 for ANOVA). There was a main effect of age, with young participants showing higher performance than older participants on average (*M* (*SD*) = 0.56 (0.22) and 0.29 (0.16), respectively). There was also a main effect of test type, with higher performance in the item test than the associative test on average (*M* (*SD*) = 0.52 (0.16) and 0.39 (0.18), respectively). Finally, and differently to the nonwords condition, there was a significant interaction between age and test type. This was because the difference between young and older adults was greater for associative than for item memory, with the older participants showing much worse performance for the associative memory test (i.e., age-related associative deficits were apparent). Once again, the pattern of results was identical using the d' data (see Tables 8 and 11).

As for the nonwords condition, floor effects were addressed by removing the worst performers from the data set based on the hits minus false alarm results of the associative test. There were one young and five older participants performing at or below chance for associative memory. Therefore, using the same logic as for the nonwords condition, the worst five participants (i.e., 21%) were removed from each age group and the above analysis repeated. As before, none of the analyses showed qualitatively different results as to when the poor associative performers were present. For a full comparison of results see Tables 8 and 11.

Like the nonwords condition, an ANCOVA was conducted on the whole dataset, to assess whether age differences in item memory could account for age differences in associative memory. Using the hits minus false alarms data, there was a significant effect of age on associative memory, even when item memory was used as a covariate, $F(1, 45) = 17.11$, $MSE = 0.03$, $p < .001$, and the covariate, item memory, was significantly related to associative memory, $F(1, 45) = 20.43$, $MSE = 0.03$, $p < .001$. Levene's test for equality of error variances was non-significant, $F(1,46) = 1.76$, *ns*, indicating that the homogeneity of variance assumption had not been violated. This shows that for words, age differences in associative memory can not be entirely accounted for by age differences in item memory. The same pattern of results was found using the d' data: There was a significant effect of age on associative memory when item memory was used as a covariate, $F(1, 57) = 14.53$, $MSE = 0.36$, $p < .001$, and the covariate, item memory, was significantly related to associative memory, $F(1, 57) = 22.33$, $MSE = 0.36$, $p < .001$.

Table 11

ANOVA Results Based on Hits Minus False Alarms (H-FA) and d' Data for Words

Effect type		H-FA		d'	
	df	*F*	*MSE*	*F*	*MSE*
Full Data					
Age	1	31.15***	0.06	29.44***	0.69
Memory test	1	54.86***	0.02	60.35***	0.19
Age x Memory test	1	15.81***	0.02	9.65**	0.19
Error	46				
Reduced Data[a]					
Age	1	38.00***	0.05	33.32***	0.62
Memory test	1	29.89***	0.01	32.09***	0.16
Age x Memory test	1	18.23***	0.01	10.16**	0.16
Error	36				

[a]Data with worst associative performers removed (5 from each age group).
** $p < .01$,*** $p < .001$.

To assess response bias, a 2 (Age: young, older) x 2 (Memory test: item, associative) repeated measures ANOVA was conducted on the ln(β) recognition data (see Figure 17 and Table 8). Unlike for the nonwords condition, there was a main effect of memory test type, $F(1, 46) = 34.05$, $MSE = 0.23$, $p < .001$, with a bias towards no/not seen before on the item test ($M = 0.69$, $SD = 0.63$), but a less biased response on the associative test ($M = 0.12$, $SD = 0.37$). There was a marginal effect of age, $F(1, 46) = 3.27$, $MSE = 0.32$, $p = .08$, but no interaction between age and test type, $F < 1$. With the worst five performers at the associative test from each age group removed to avoid floor effects, the main effect of memory test remained, *F*(1,

36) = 23.88, *MSE* = 0.26, *p* <.001 (item test *M* = 0.71, *SD* = 0.68, associative test *M* = 0.14, *SD* = 0.39). Also the main effect of age remained non significant, *F(*1, 36) = 2.59, *ns*, as well as the age by test type interaction, *F* < 1.

Three-way analysis. A 2 (Age: young, older) x 2 (Memory test: item, associative) x 2 (Condition: words, nonwords) repeated measures ANOVA was conducted on the data from both conditions (see Tables 8 and 12). For the nonword data, only the three repetitions young data were used in this analysis. Age and condition were between-subjects factors and test type was a within-subjects factor. Using hits minus false alarms, there was a main effect of age, with young participants showing higher performance than older participants on average (*M* (*SD*) = 0.58 (0.21) and 0.34 (0.20), respectively). There was a main effect of condition, with higher performance overall in the nonwords condition than the words condition (*M* (*SD*) = 0.49 (0.23) and 0.42 (0.19), respectively). There was also a main effect of test type, with performance overall higher in the item test than the associative (*M* (*SD*) = 0.60 (0.16) and 0.32 (0.26), respectively).

There was an interaction between test type and age. There was also an interaction between test type and condition. The difference in item memory performance between the two conditions was greater than the difference in associative memory performance (the nonwords condition had a greater item memory performance). This was probably because nonwords had no preexisting concepts in memory; all that was needed to correctly recognise a seen-before non-word item was a sense of any familiarity. With words, previously unseen words already existed in memory and were already familiar in some way. There was no interaction between age and condition. The crucial interaction between age, test type, and condition was marginal (*p* = .08). This indicates that the difference between item and associative memory performance across age tended to be larger in the words condition than the nonwords condition. Using d' values, the three-way interaction became significant, and the main effect of condition became marginal (*p* = .06). Also, the interaction between test type and age was no longer significant. Other d' statistics showed the same numerical trends as the hits minus false alarms data (see Tables 8 and 12).

The above analysis was repeated with the worst performers from each age group removed – the 12 worst for the nonwords condition and the five worst for the words condition for both young and older groups. For the hits minus false alarm analysis, there remained main effects of age, test type and condition. All of the interactions remained present and the new analysis now showed a significant three-way interaction between age, test type and condition.[21] This is an important result as it indicates that the absence of age-related associative deficits with nonwords is significantly different to the age-related associative deficits found with words. Using d' with the reduced data showed qualitatively identical pattern of results to hits minus false alarms data except that the test type by age interaction was non-significant. Unlike the previous d' analysis, the main effect of condition was now significant and congruent with the hits minus false alarms data. For a full comparison of results see Tables 8 and 12.

[21] The analysis was also conducted with the same proportion of participants excluded from each condition (the worst 40% of young and older performers from the words and nonwords conditions). All the main effects and interactions were the same as above, including the crucial significant three-way interaction between age, test type and condition.

Table 12

ANOVA Results Based on Hits Minus False Alarms (H-FA) and d' Data

Effect type		H-FA		d'	
	df	*F*	*MSE*	*F*	*MSE*
Full Data					
Age	1	49.27***	0.07	49.56***	0.74
Memory test	1	141.73***	0.03	177.11***	0.30
Condition	1	4.18*	0.07	3.56[*]	0.74
Age x Memory test	1	6.97*	0.03	2.57	0.30
Memory test x Condition	1	13.29***	0.03	16.70***	0.30
Age x Condition	1	0.43	0.07	0.62	0.74
Memory test x Age x Condition	1	3.22[*]	0.03	4.35*	0.30
Error	104				
Reduced Data[a]					
Age	1	56.07***	0.05	51.45***	0.60
Memory test	1	85.77***	0.01	110.31***	0.16
Condition	1	14.81***	0.05	10.03**	0.60
Age x Memory test	1	9.48**	0.01	2.72	0.16
Memory test x Condition	1	4.17*	0.01	6.75*	0.16
Age x Condition	1	1.10	0.05	1.02	0.60
Memory test x Age x Condition	1	6.55*	0.01	7.83**	0.16
Error	70				

[a]Data with worst associative performers removed (12 from each age group excluded for the nonwords condition and 5 from each age group for the words condition).
[*] $p < .1$, * $p < .05$, ** $p < .01$,*** $p < .001$.

To further assess response bias, a 2 (Age: young, older) x 2 (Memory test: item, associative) x 2 (Condition: words, nonwords) repeated measures ANOVA was conducted on the ln(β) data from both conditions (see Figure 17 and Table 8). There was no main effect of age, $F < 1$. There was a main effect of memory test type, $F(1, 104) = 23.94$, $MSE = 0.29$, $p < .001$. This is because there was a bias towards no/not seen before responses for the item test ($M = 0.40$, $SD = 0.71$) but not for the associative test ($M = 0.04$, $SD = 0.40$). There was also a main effect of condition, $F(1, 104) = 17.57$, $MSE = 0.41$, $p < .001$; there was a bias towards no/not seen before responses in the words condition ($M = 0.41$, $SD = 0.50$), but not for the nonwords condition ($M = 0.04$, $SD = 0.61$). There was an interaction between memory test type and condition, $F(1, 104) = 8.59$, $MSE = 0.29$, $p < .01$. This is because for the nonwords condition, participants showed only a small bias for item and associative tests, but for the words condition there was a large bias towards no/not seen before responses in the item test but a less biased response in the associative test. There was also a significant interaction between age and condition, $F(1, 104) = 5.13$, $MSE = 0.41$, $p < .05$. Young participants showed a larger difference in response bias between conditions ($M = 0.51$ and -0.06 for words and nonwords conditions, respectively) than older participants ($M = 0.30$ and 0.13 for words and nonwords conditions, respectively). There was no interaction between test type and age, $F(1, 104) = 1.63$, *ns*. The three-way interaction between age, condition and memory test type was marginally significant, $F(1, 104) = 3.84$, $MSE = 0.29$, $p = .05$.

As with earlier analyses, the worst performers on the associative test were removed and the response bias analysis repeated (see Table and 8 for means). For the words condition, five young and five older participants were excluded, and for the nonwords condition, 12 young and 12 older participants were excluded. There remained no main effect of age, $F < 1$. There also remained main effects of memory test type, $F(1, 70) = 18.69$, $MSE = 0.27$, $p < .001$, and of condition, $F(1, 70) = 9.05$, $MSE = 0.48$, $p < .01$. The interaction between memory test type and condition remained, $F(1, 70) = 5.65$, $MSE = 0.27$, $p < .05$. The interaction between age and condition lost significance and became marginal, $F(1, 70) = 3.06$, $MSE = 0.48$, $p = .09$, and the test type by age interaction remained non-significant, $F(1,70) = 1.10$, *ns*. Finally, the three-way interaction (Age x Test Type x Condition), which was previously marginal, completely disappeared, $F < 1$.

Reaction times. For the nonwords condition, data were also gathered for reaction times when responding to each trial of the item and associative recognition tests. There were four categories of response: Hit (H) – Correctly making a positive response to a seen before stimulus; Correct rejection (CR) – Correctly making a negative response to a previously unseen stimulus; Miss (M) – Incorrectly making a negative response to a seen before stimulus; and False alarm (FA) – Incorrectly making a positive response to a previously unseen stimulus. For each participant, the reaction time to the first trial of each test (item and associative) was excluded. This is because participants were often familiarising themselves with the buttons on the first trial and reacted much slower than normal.

Figure 18 shows the mean reaction times for each response category. To summarise the data, item tests generally showed faster reactions than associative tests. Young participants were generally faster than older participants; however, age differences were greater for the two repetition young group than for the three repetition young group compared to the older group. In terms of age-related associative deficits, age differences were generally larger for the associative test than for the item test. Additionally, correct responses were generally faster than incorrect responses for both endorsements (hits vs. false alarms) and rejections (correct rejections vs. misses). This indicated that there was no speed-accuracy trade-off because accuracy was higher for the faster responses.

A 2 (Age: young 3 repetitions, older) x 2 (Test type: item, associative) x 4 (Reaction category: H, CR, M, FA) repeated measures ANOVA was conducted on the reaction time data (see Figure 18[22]). Only 12 young and 20 older participants were included in this analysis because many participants did not make a response for each of the four reaction categories for each test type.

There was a main effect of age, $F(1, 30) = 9.49$, $MSE = 5.86 \times 10^6$, $p < .01$, with young participants reacting quicker than older participants on average (M (SD)= 2259 (856) ms and 3222 (854) ms, respectively). There was a main effect of test type, $F(1, 30) = 29.19$, $MSE = 3.42 \times 10^6$, $p < .001$, with reaction times quicker for the item than the associative tests on average (M (SD) = 2096 (645) ms and 3385 (1431) ms, respectively). There was also a main effect of category, $F(3, 90) = 6.20$, $MSE = 1.03 \times 10^6$, $p = .001$, with hit responses being generally quicker than the other three response types. None of the interactions was significant, although the test type by age interaction was marginal, $F(1, 30) = 3.72$, $MSE = 3.42 \times 10^6$, $p = .06$. The reaction time difference between young and older participants tended to be larger for associative than for item tests as the older participants appeared to be extra slow compared to young participants at reacting to the associative trials.

---

[22] Note that Figure 18 uses more data than the repeated measures ANOVA. Each bar in the figure was produced with the average response for that category. However the ANOVA is constrained to only include participants who made a response for every category.

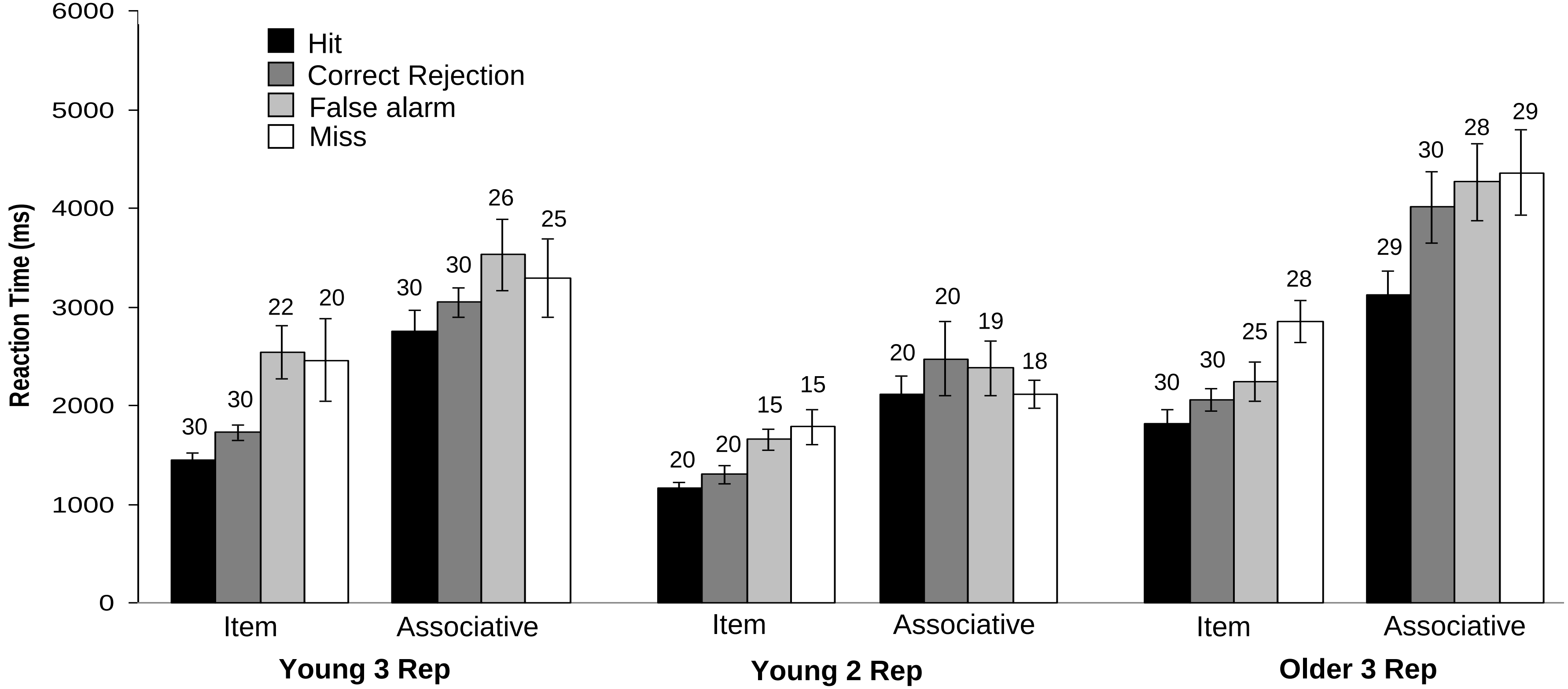

Figure 18. Reaction times for each response category of the nonwords condition: Hit, correct rejection, false alarm and miss. Data are shown for both item and associative test types, for young (3 and 2 repetitions of memory set) and for older participants. Error bars are ± 1 *SE*. Numbers above bars show the number of participants who made a response that contributed to each statistic out of 30, 20, and 30 participants respectively.

The 2 x 2 x 4 repeated measures ANOVA was computed for the two repetition young group (see Figure 18). With this analysis, Mauchly's test indicated that there was a violation of sphericity for the test type by reaction category interaction, $\chi^2(5) = 13.13$, $p < .05$; therefore the test type by reaction category degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .78$). Ten young and 20 older participants made responses for every category and were included in the analysis.

The pattern of results was very similar to the three repetition young group. There was a main effect of age, $F(1, 28) = 15.73$, $MSE = 5.85 \times 10^6$, $p < .001$, with young participants reacting quicker than older participants on average (*M* (*SD*) = 1908 (854) ms and 3222 (854) ms, respectively). There was a main effect of test type, $F(1, 28) = 25.35$, $MSE = 3.55 \times 10^6$, $p < .001$, with reaction times quicker for the item than the associative tests on average (*M* (*SD*) = 1916 (635) ms and 3215 (1495) ms, respectively). There was also a main effect of category, $F(3, 84) = 3.60$, $MSE = 1.08 \times 10^6$, $p = .05$, with hit responses being generally quicker than the other three response types. None of the interactions was significant, although the test type by age interaction was marginal, $F(1, 28) = 3.05$, $MSE = 3.55 \times 10^6$, $p = .09$.

Discussion

The current study provides evidence that young and older adults differentially remember familiar and unfamiliar stimuli. This difference in familiarity between words and nonwords can be seen as a difference in the amount of preexisting knowledge related to the stimuli. The words condition produced typical age-related associative deficits, with age differences greater for associative memory than for item memory.

However, when unfamiliar nonwords were used with the same types of memory tests, age differences were constant between associative and item memory. It is therefore suggested that age-related associative deficits with words are observed because the item test is for concepts supported by preexisting knowledge while the associative test is for completely new concepts: Age differences on the item test for words are reduced compared to associative memory age differences because older participants can use preexisting knowledge of the words to support memory formation. This view is shared by MacKay and Burke (1990), who proposed a commitment learning principle whereby age differences in memory formation are smaller when fewer new connections are required. In their chapter, they discussed a variety of evidence showing that older adults produce smaller age differences in word memory tasks when they can use preexisting knowledge (referred to as engrainment learning) to support the encoding of information.

One issue regarding this explanation is that for the nonwords item test, the age difference was similar to that of the words item test despite a lack of preexisting knowledge for nonwords. It could therefore be argued that older adults were not benefiting from preexisting knowledge in the words item test as we would expect to see a larger age deficit in item memory when using the novel nonwords. Note, however, that the experimental design was not suited to such direct comparisons between the conditions; this is explored in more detail later in the discussion.

Alternatively, the data can be viewed in terms of dual process models of memory, which could account for the full pattern of results including the above issue. In dual process models, familiarity and recollection are considered as different memory processes (see Chapter 1 for review). Familiarity is seen as more automatic, providing a sense that a given stimulus has been encountered before following a recognition probe, whereas recollection is seen as controlled conscious retrieval of specific episodic experiences. Dual process models commonly assume that during recognition memory tasks, when the level of familiarity/unfamiliarity in memory is ambiguous, a further recollection based memory search is required before a response can be made (Yonelinas, 2002). Familiarity based processes are considered sufficient to complete an item test whereas associative tests are more reliant on recollection based processes (Healy et al., 2005; Old & Naveh-Benjamin, 2008a).[23] This could explain why age deficits are often smaller for item tests than for associative tests because age deficits are typically smaller for recognition (familiarity based memory) than for recall (recollection based memory) (e.g., Craik & McDowd, 1987; Schonfield & Robertson, 1966).

It may be the case that the nonwords condition of the current experiment differs from the words condition in terms of its relative reliance on familiarity vs. recollective processes. In particular, it seems more likely that recollection would contribute to associative memory for words than for nonwords because of the former's greater preexisting semantic knowledge. For example, unrelated word pairs such as contest-dancer may nevertheless evoke a specific mental image at encoding that can be recollected at test, whereas nonword pairs such as bligma-slanquil are unlikely to do so. Thus age-related associative deficits would be larger with words than with nonwords. As can be seen in Figures 15 and 16, age differences were largest for associative memory with words but were smaller for item memory with words and both item and

---

[23] Note that the associative test was specifically designed to minimise reliance on familiarity in the recognition process: The associative recognition test always showed familiar (seen before) items in pairs and the test was to establish if the *combination* of those items was intact or recombined. Therefore intact and recombined test pairs would evoke similar levels of familiarity which, in line with dual process models, would require recollection processes to produce a correct response.

associative memory with nonwords. It may be that all of the tests are based mainly on familiarity except for the words associative test, although the reaction time data suggest that recollection may be involved in the nonwords associative test.

For reaction times in the nonwords condition, it was generally found that young participants were faster than older participants for all response types (hits, correct rejections, false alarms[24] and misses) over both item and associative tests. This is consistent with the general slowing hypothesis (Salthouse, 1991), which suggests that cognitive processes become slower in older adults, hindering their ability to perform as well as young adults. Hit responses were faster than the other response categories for both age groups and both test types. This can also be accounted for by dual process models of memory where familiarity and recollection are considered as different processes (see Yonelinas, 2002, for review). It is likely that seen-before stimuli elicited a sense of familiarity which enabled a quicker response. Likewise, correct rejection responses were faster on the whole than false alarms and misses, accountable for by a sense of unfamiliarity. Dual process models generally indicate that when the level of familiarity/unfamiliarity in memory is ambiguous, a further recollection based memory search is required before a response can be made. A range of dual process models of memory agree that familiarity is faster than recollection (see Yonelinas, 2002, for review); therefore, this explanation fits the pattern of the reaction time data, because responses that were more based on familiarity tended to be quicker. Furthermore, for the associative tests, reaction times were generally slower than the item tests. In the associative tests all items were previously seen so it is unlikely that familiarity based

---

[24] Young participants who saw three repetitions of the memory set had slightly slower item test false alarms than older participants.

mechanisms would be entirely sufficient to produce a confident response. It is possible that recollection based mechanisms were used to determine if the nonword pairs were originally seen together.

Reaction time data were not available for the words condition. This meant that it was not possible to test the idea expressed above that all conditions relied mainly on familiarity except the words associative test. If this was the case it would be expected that reaction times for the words associative test would be the slowest of all tests because responses were more likely to be based on recollective processes. It seems that such data are not usually gathered as none of the papers that showed associative deficits also included reaction times in their results. There is evidence in the literature, however, showing that when the item test is recall based, age-related associative deficits are not present (see Old & Naveh-Benjamin, 2008a, for review). This is presumably because familiarity cannot be used to complete the item memory test and item memory age differences are increased to a similar level as associative memory age differences.

A related point to note is that young adults are generally more likely than older adults to implement encoding strategies (e.g., Luszcz et al., 1990; Witte et al., 1990). However, it is possible that young adults are unable to implement an encoding strategy for associating nonwords as easily as they are able to do so for words. This may equate young and older adults on the implementation of encoding strategies (and possibly the balance of familiarity/recollection) because neither group is likely to incorporate an encoding strategy with nonwords. Thus, the nonwords condition may actually provide a ‘purer’ test of the associative deficit hypothesis.

In the literature, it is generally found that memory is improved in areas of expertise, where there are likely to be more familiar/preexisting concepts to aid cognition (e.g., Chase & Simon, 1973; Gobet & Simon, 1998). In the context of ageing, expertise has been found to benefit both young and older age groups in memory tasks (see Zacks et al., 2000, for review), although contrary to the hypothesis expressed here, the presence of expertise (preexisting knowledge of items) did not reduce age differences in prior research. A further point to consider is that although expertise may benefit young and older participants equally, the level of expertise with words was different between the two age groups. It was shown earlier that the older participants in the nonwords condition had superior vocabulary ability to the young participants. Although these data were not available for the words condition, the general finding is that older participants have superior vocabulary to young participants (see Verhaeghen, 2003, for review). The greater vocabulary of older participants may have helped them memorise the words – a thorough investigation by Hultsch, Hertzog, and Dixon (1990) showed that word memory ability correlated highly with vocabulary for both young and older age groups. This evidence suggests that increased preexisting knowledge of vocabulary in older participants compared to young participants may have partly aided their recognition of words in the word condition item test and attenuated age differences for item memory.

Young and older participants showed similar patterns of response bias. However, the item test for words showed a bias towards no/not-seen-before whilst the other tests were relatively unbiased. This is noteworthy because it was the only test conducted that involved the recognition of concepts that were preexisting in memory prior to the experiment. Participants may therefore have been less confident (and hence more cautious) in responding yes/seen-before in the words item memory test because preexisting knowledge was available for both old and new words at test.

The nonwords condition was completed with two groups of young participants: one group received three repetitions of the memory set exactly the same as the older group, and the other group saw two repetitions of the memory set. The age differences were successfully removed when comparing the two-repetition young group to the older group, indicating that the lack of an age by test interaction was not due to task difficulty. Previous research with words has shown that age-related associative deficits are still present when young and older participants have similar performance on item memory (Naveh-Benjamin, Guez, & Shulman, 2004; Naveh-Benjamin, Hussain et al., 2003). In those experiments, task difficulty was increased for young participants by dividing their attention during encoding (the effects of divided attention on associative memory are explored in more detail in Chapter 8). As with nonwords, increasing the task difficulty for young participants did not differentially affect their item and associative memory performance (i.e., it did not produce associative deficits). Naveh-Benjamin et al. (2004) argue that age-related associative deficits are (1) unlikely to be due to reduced cognitive resources in older adults but rather they are (2) probably due to a specific deficit in associative memory. The nonwords results agree with the first part of this argument, as task difficulty did not differentially affect the item and associative tests – changing the task difficulty did not affect one test more than the other. Therefore, differences in cognitive resources cannot account for greater age differences in associative tests compared to item tests. However, for the second part of the argument, the nonwords results suggest that typical associative deficits in older participants may be due to a deficit in *novel* rather than associative memory.

A potential problem with the current experimental design is that in the words condition, only one presentation of the memory set occurred but with nonwords there were three. This is problematic as it invalidates any direct comparison of overall performance levels between the two conditions; however, the comparison of age-related associative deficits remains valid. Overman and Becker (2009) observed that repetition of a memory set benefited both young and older participants for item memory but just young participants for associative memory. Similar studies of associative memory have shown that older adults benefit equally to young adults from repetition (Kornell, Castel, Eich, & Bjork, 2010) or less than young adults from repetition (Light et al., 2004). Several studies have also shown that young and older adults benefit equally from spacing compared to massed presentation of associative memory stimuli (e.g., Balota, Duchek, Sergent-Marshall, & Roediger, 2006; Benjamin & Craik, 2001; Logan & Balota, 2008). Therefore, the (spaced) repetition of the memory set in the current nonwords condition should have either increased age-related associative deficits or left them unchanged. This literature suggests that repetition is unlikely to account for the absence of age-related associative deficits with nonwords. In addition, the effect of repetition was analysed by comparing the nonwords data from the young groups with two vs. three repetitions. There was a significant benefit from the additional repetition but this was equivalent for item and associative tests as indicated by no interaction between repetitions and test type, $F < 1$. Overall there are many studies demonstrating age-related associative deficits with different experimental parameters and there is no evidence to suggest that varying the numbers of repetitions, presentation rates, or memory set sizes would eliminate age-related associative deficits.

Finally, the present results raise an interesting issue of how to categorise an ‘item’. It could be argued that the nonword items required associative memory because participants had to associate the first syllable with the second syllable to remember each of them. Thus, the pattern of results could be explained in terms of associative deficits – the absence of age-related associative deficits with nonwords could be due to both ‘item’ and associative memory being ultimately associative in this case. However, such an explanation implies a reconsideration of what is meant by an item. Perhaps an item is not a unit of information that is presented and remembered but rather a unit of information that can be represented by a single preexisting concept. Therefore, understanding age-related associative deficits could lead to insight into how preexisting knowledge is used to encode memory and why this process is more robust against the ageing process.

Experiment 4 indicates that it is important to consider how item and associative memory are defined in measures of age-related associative deficits. Item memory may be defined as a preexisting concept reactivated and associative memory may be defined as new links between existing information in memory. This implies that age-related associative deficits are due to age deficits in forming new/novel links in memory. Despite many years of research, it is difficult to create an experiment that tests item and associative memory in an entirely comparable way and the distinction between the two memory types remains incomplete. The following chapter further explores the potentially important role of preexisting knowledge in associative memory.

Chapter 6: Integrative and Semantic Relations Equally Alleviate Age-Related Associative Memory Deficits

In the previous chapter, it was demonstrated that age-related associative deficits are reduced when memory stimuli are novel (nonwords) compared to when the stimuli are familiar (words). It was hypothesised that this pattern of results could be explained if older adults can make use of preexisting knowledge to reduce age deficits in memory. The current chapter aimed to manipulate preexisting knowledge between to-be-associated items in order to directly assess the use of preexisting structures in associative memory formation.

Research into age-related associative deficits has attempted to establish factors that can alleviate this memory deficit. One such factor is the semantic relatedness between to-be-associated items. Items are *semantically related* if they belong in the same category, such as *shirt* and *sock*, or are otherwise featurally similar, such as *apple* and *ball*. Naveh-Benjamin (2000, Exp. 4), Naveh-Benjamin, Hussain et al. (2003, Exp. 2) and Naveh-Benjamin et al. (2005) showed a reduction in age differences for associative memory with semantically related word pairs (e.g., *shirt* and *sock*) compared to unrelated word pairs (e.g., *shirt* and *apple*). Therefore, older adults are able to use semantic relations to enhance their associative memory performance relative to young adults. This finding suggests that older adults' associative memory deficit may be specific to *new* associations; older adults' memory for *preexisting* associations appears to be relatively unimpaired. Indeed, MacKay and Burke (1990) and Naveh-Benjamin, Hussain et al. (2003) both suggested that age-related memory deficits increase on tasks that require more novel associations. In the previous chapter, it was shown that several recent studies support this claim (e.g., Castel, 2005, 2007; Patterson et al., 2009).

Semantic relations may alleviate older adults' associative deficits in multiple ways. First, semantic relations may allow older adults to make use of overlapping neural representations: The co-activation of features shared by semantically related items may strengthen the associative memory representation that links them (MacKay & Burke, 1990). In contrast, because semantically unrelated items have more distinct neural representations, that lack of co-activation would produce weaker associative memory representations. Second, older adults may use semantic relations to initiate encoding and retrieval strategies during memory tasks. A consistent finding in the literature is that older adults are less likely than young adults to implement an encoding strategy (e.g., Luszcz et al., 1990; Witte et al., 1990). Therefore, semantically related word pairs could show smaller age deficits than unrelated word pairs because with semantic relations young and older adults are better equated in their use of encoding and retrieval strategies. That is, with semantically related word pairs, older adults may more easily adopt a strategy to aid the memory process, whereas with unrelated word pairs older adults may not produce encoding and retrieval strategies as well as young adults.

Prior studies have successfully reduced older adults' associative memory deficits by introducing preexisting relations between items. The current study also aims to reduce this memory deficit but without recourse to preexisting relations. Specifically, the study examined the age-related associative deficit with three different types of word pairs: integrative word pairs, semantically related word pairs and unrelated word pairs. The novel element of this study was the use of integrative word pairs. These are word pairs where the two words can be linked together to produce a coherent phrase (e.g., *horse-doctor*, *plastic-toy*). Ultimately, any word pair in which the first word modifies the second word involves integration. Although this includes simple adjective-noun pairs such as *red apple*, it also includes noun-noun pairs such as *thesis idea,* which are more common among studies of memory. Integrative relations entail a modifier (i.e., first word) that specifies a subclass of the head noun (i.e., second word). For example, a *thesis idea* is a specific type of idea, and a *trick rabbit* is a specific type of rabbit that differs in important respects from the more general class of rabbits (e.g., Glucksberg & Estes, 2000; Springer & Murphy, 1992).

Notably, many words can be integrated easily despite being semantically dissimilar, unassociated, and unfamiliar as a phrase (for review see Estes, Golonka, & Jones, 2011). *Monkey foot*, for instance, is easily understood despite the fact that *monkey* and *foot* are dissimilar and do not occur together frequently in language. Such integrative word pairs lack preexisting relations: They are from different semantic categories, they share few features (if any), they are rarely spoken or written together, and they rarely occur together in a free association task (Estes & Jones, 2009). This novel aspect of integration allowed the testing of older adults' processing of and memory for integrative word pairs that have few preexisting relations between them (like the unrelated word pairs) but could very easily be encoded together (like the semantic word pairs). If integrative word pairs produced small age-related associative deficits like semantic word pairs, then this would indicate that ease of encoding/retrieval can reduce associative deficits. Alternatively, if integrative word pairs produced larger age-related associative deficits than semantic word pairs (like unrelated word pairs), then this would indicate that preexisting relations are a key factor that reduces associative deficits.

Integrative Priming and Memory

Integrative relations facilitate processing of words. Estes and Jones (2009) demonstrated integrative priming in young adults. In their Experiment 2, integrative priming was directly compared to semantic priming. Participants were presented with trials where a prime was followed by a target. They completed a lexical decision task where they had to decide if each target was a word or a nonword. Prime-target pairs were either integrative or semantic word pairs. There was also a baseline condition where the prime word was replaced by a row of asterisks. Both integrative and semantic primes facilitated the lexical decisions as responses were significantly faster than responses to the baseline condition. There was also no significant difference between the magnitudes of integrative and semantic priming.

The integrative priming effect is interesting because the faster response times following integrative primes cannot be explained by pre-processing of the prime before the target onset (Estes & Jones, 2009). For example, with the semantic prime-target pair *fox-dog,* semantic elaboration of the features of a fox will act before the target *dog* appears and therefore the response to *dog* is facilitated. However, with integrative pairs (e.g., *apartment-dog*), the prime is unlikely to activate the target as the two words are initially unrelated. This means that integrative priming processes occur *after* viewing the target. In terms of the current study, integrative word pairs are important for discriminating between memory processes that occur only upon encoding and retrieval (i.e., integrative pairs) and those that may also rely on preexisting relations (i.e., semantic pairs). However, because integrative priming has not yet been demonstrated in older adults, Experiment 5 of this thesis replicated this effect in older adults. It is well established in the literature that older adults demonstrate semantic priming to at least the same extent as young adults and possibly to a greater extent (e.g., Laver, 2009; see Laver & Burke, 1993; Myerson, Ferraro, Hale, & Lima, 1992, for reviews). In the first experiment, integrative and semantic priming were compared in older adults to establish if older adults produce an integrative priming effect.

Integrative relations also facilitate memory. Jones, Estes, and Marsh (2008) argued that conceptual integration may elicit elaboration during encoding and may act as a contextual cue during retrieval. In support of this argument, Jones and colleagues reported two experiments in which integrative relations affected memory in young adults. First they presented word pairs that were significantly easier to integrate in one order (e.g., *horse doctor*) than in the reverse order (e.g., *doctor horse*). They subsequently presented those individual words in a surprise recognition memory test. They found that the words were more reliably recognised if they had been studied in their more easily integrated order (i.e., *horse doctor*) than if they had been studied in their less integratable order (i.e., *doctor horse*). In another experiment, Jones and colleagues showed that a given item was more reliably recognised at test when it instantiated the same integrative relation at study than when it instantiated a different relation. For example, the item *cookie* was better recognised in *cookie plate* when it had been studied as *cookie jar* than when it was studied as *cookie crumb*. Because *cookie jar* and *cookie plate* both instantiate a containment relation (i.e., Y contains X), the target item was more reliably recognised. Thus, both of these experiments indicate that integrative relations facilitate item memory.

Demonstrating integrative priming in older adults (Experiment 5) would validate the use of integrative word pairs in a memory test with older adults (Experiment 6), in that the observation of integrative priming among older adults would justify the assumption that encoding and retrieval of integrative word pairs is relatively easy for older participants. Given that integrative relations facilitate word processing (Estes & Jones, 2009) and item memory (Jones et al., 2008) in young adults, we hypothesised that integrative relations might similarly facilitate word processing and associative memory in older adults, despite the lack of preexisting relations between the words.

Experiment 5a

To establish if older adults produce the same integrative priming effect as young adults, the exact experimental procedure for the 500 ms stimulus onset asynchrony (SOA) part of Estes and Jones' (2009) Experiment 2 was replicated. This would allow comparison of the current experiment to the data gathered by Estes and Jones (2009) who tested young adults and produced measures of integrative and semantic priming.

Method

Materials. The integrative and semantic prime-target word pairs (see Appendix 4) were acquired from Estes and Jones (2009) where the stimuli were selected based on results from pretesting: Twenty-four participants rated the stimuli based on a seven-point integratability scale (1 = *not linked* to 7 = *tightly linked*) and on a seven-point semantic similarity scale (1 = *not similar* to 7 = *very similar*). In addition, integrative and semantic pairs were chosen to have low levels of both forward (i.e., *prime-target*) and backward (i.e., *target-prime*) association probabilities taken from the University of South Florida free association norms (Nelson, McEvoy, & Schreiber, 2004); see Table 13 for a summary and Estes and Jones (2009) for further details. For a given target there was an integrative and a semantic prime. Integrative primes were selected to have a high rating of integratability and a low rating of semantic similarity to the target. Semantic primes were selected to have a high rating of semantic similarity and a low rating of integratability to the target.

Table 13

Integratability Ratings, Semantic Similarity Ratings and Forward and Backward Association Probabilities for the Materials Used in Experiments 5 and 6

	Integratability		Semantic Similarity		Association			
					Forward		Backward	
Pair type	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Integrative	5.41	0.85	2.14	0.87	0.02	0.03	0.00	0.01
Semantic	3.00	0.74	4.68	0.77	0.02	0.02	0.01	0.02

Note. Adapted from “Integrative Priming Occurs Rapidly and Uncontrollably During Lexical Processing,” by Estes and Jones, 2009, *Journal of Experimental Psychology: General*, *138*, p. 116. Copyright 2009 by the American Psychological Association.

In total there were 45 target words, each corresponding to one of 45 integrative primes and one of 45 semantic primes (e.g., for the target *foot*, the integrative prime was *monkey* and the semantic prime was *paw*). For the lexical decision task, there were 45 nonword targets which were also those employed by Estes and Jones (2009). Three separate lists were produced for counterbalancing, each containing 90 prime-target pairs.

For each list there were 15 integrative primes, 15 semantic primes, and 15 baseline primes (the baseline primes were a row of eight asterisks). The remaining 45 pairs consisted of nonword targets, 15 with asterisk primes and 30 with word primes. The lists were counterbalanced so that for a given real word target, one list would contain an integrative prime, one a semantic prime and one a baseline prime. In this way, no two counterbalanced lists had any of the same prime-target pairs: There were six participants in each counterbalancing condition who saw different combinations of prime-target pairs.

Participants. Thirty older adults (21 female), aged 58-88 years (*M* = 74.8, *SD* = 7.5), took part in the experiment. They were recruited from the University of Warwick Age and Memory Study volunteer panel that was populated by local advertisements; they were offered no financial incentives for participation.

To assess cognitive functioning, participants completed the Digit Symbol Substitution task (Wechsler, 1981) as a measure of perceptual-motor or processing speed (*M* = 44.23, *SD* = 10.69). They also completed the multiple choice part of the Mill Hill vocabulary test (Raven et al., 1988) as a measure of crystallised intelligence (*M* = 22.53, *SD* = 3.63).

Procedure. Participants were informed that they were going to be shown words on a computer screen and that their task was to identify whether or not those words were real or nonwords. Participants were informed that they would see a red word or row of asterisks before each target word and that they were not to respond to it but to base their word/nonword judgement on the white word that followed it. For each trial, a fixation cross appeared on a blank black computer screen in red for 500 ms. This was then immediately followed by a prime word/asterisks in red for 100 ms. There was then a 400-ms delay with a blank black screen followed by the target word in white. This corresponds to a 500-ms SOA as in Estes and Jones' (2009) Experiment 2. Once the target appeared on the screen, participants were required to press the 'j' key on the keyboard if the target was a word and to press the 'f' key if the target was a nonword. After a response was made, the screen displayed the instruction 'Press space when ready' in white and participants needed to press the space bar to activate the next trial.

Words were presented in a lower case font size of 40 pt with a height corresponding to approximately 1.4^{o} viewing angle at a distance of 60 cm. Participants were required to keep their index fingers ready on the 'f' and 'j' keys and to press space between each trial with their thumb. Before the main task, participants completed a practice test with 10 trials of mixed prime type using separate stimuli to the main study. If a participant did not follow instructions properly, they were encouraged to practice again by the experimenter. Participants were also given a reminder sheet that identified which button was which.

Results and Discussion

Only correct responses for the lexical decision task were used to formulate response speed averages for each trial type. For each prime type, the average reaction time of correct responses to word targets was calculated on an individual basis for each participant; responses falling outside of 2.5 standard deviations from each average were excluded as outliers. The data from the 500-ms SOA condition of Estes and Jones (2009) (42 young participants) were used to compare young and older adults. The exact ages of these participants were not available; however they were all undergraduates at the University of Georgia (USA) so were presumed to be young adults.

Firstly, to assess any effects of counterbalancing, a 3 (Prime type: integrative, semantic, baseline) x 3 (List type: 3 levels of counterbalancing between subjects) repeated measures ANOVA was conducted on the reaction time data for older adults.[25] There was no main effect of list type, $F < 1$, and no interaction between prime type and list type, $F(4, 54) = 1.16$, $MSE = 8{,}857$, *ns*. This shows that the different lists used did not affect the data pattern so the following analyses were conducted without the list type factor.

To assess the data as a whole, a 3 (Prime type: integrative, semantic, baseline) x 2 (Age: young, older) repeated measures ANOVA was conducted on the reaction time data (see Figure 19A). There was no main effect of prime type, $F(2, 140) = 1.77$, $MSE = 5{,}359$, *ns*, which initially indicated no priming in the data. There was a main effect of age, $F(1, 70) = 31.02$, $MSE = 9.40 \times 10^4$, $p < .001$. There was also a significant interaction between prime type and age, $F(2, 140) = 3.07$, $MSE = 5{,}359$, $p < .05$; it can be seen in Figure 19B that the young adults produced positive priming and the older adults showed no priming.

To follow up the interaction, the priming effects were measured for integrative and semantic priming for young and older adults by conducting pairwise *t*-tests on the integrative vs. baseline and the semantic vs. baseline reaction times. For young adults, there was significant integrative and semantic priming, $t(41) = 3.57$, $p = .001$, and $t(41)$

[25] The counterbalancing data were not available for Estes and Jones (2009) so this analysis was only conducted on data from the current experiment.

= 4.35, $p < .001$, respectively. For the older adults, there was no integrative or semantic priming, $t(29) = -0.25$, *ns*, and $t(29) = -0.31$, *ns*, respectively.

The percentage of correct responses was also analysed. All mean percentages of correct responses for young and older adults for every prime type were greater than 98% correct. A 3 (Prime type: integrative, semantic, baseline) x 2 (Age: young, older) repeated measures ANOVA was conducted on the proportion of correct responses. There was no main effect of prime type, $F < 1$, or age, $F(1, 70) = 1.76$, $MSE = 0.001$, *ns*, and no interaction between prime type and age, $F < 1$. This is because of ceiling performance: the lexical decision task was relatively easy to do as none of the word targets used was unusual.

To summarise, the young participants demonstrated a priming effect but the older participants did not. As semantic priming in older adults is a reliable phenomenon, it would be unwise to draw any conclusions about integrative priming in older adults from these data when semantic priming could not be replicated. This was most likely due to the salience of the prime; upon debriefing, many of the older adults said they did not really notice the prime. The prime was only displayed for 100 ms. Examining those studies of semantic priming with older adults reviewed by Laver and Burke (1993) that provided information regarding prime duration and SOA, all used longer prime durations and most used longer SOAs than the current study. Thus, in Experiment 5b the salience of the prime was increased in order to produce a reliable priming effect in older adults.

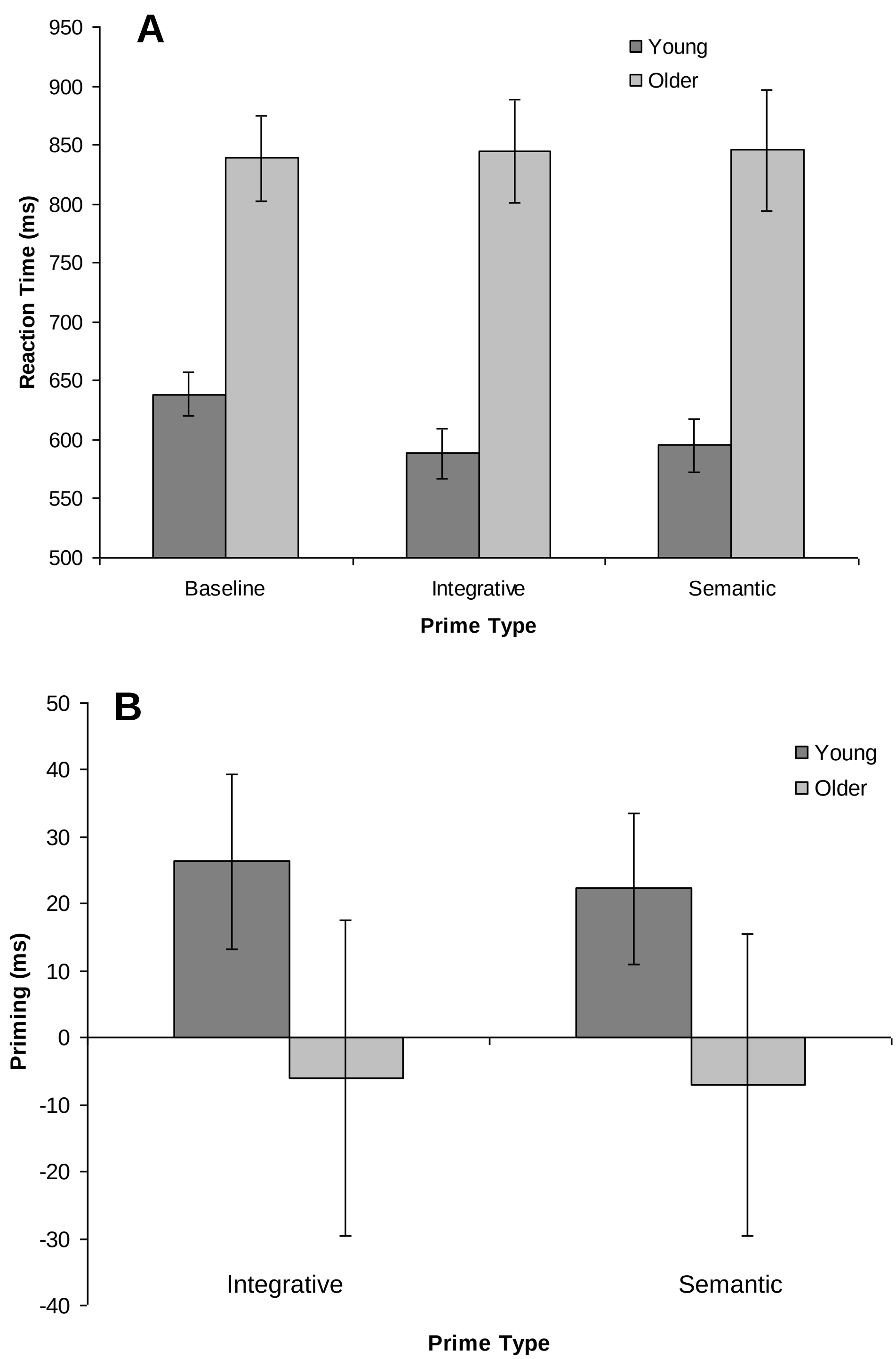

Figure 19. A: Mean correct reaction times to targets following baseline, integrative and semantic primes. B: Integrative and semantic priming in young and older adults. Young data are from Estes and Jones (2009). Error bars are ±1 *SE*.

Experiment 5b

Before testing a large group of participants with a new experimental set up, two pilot studies were conducted to establish the appropriate parameters to produce a priming effect in older adults.

Pilot 1. Initially a small group of eight older participants (4 female) aged 66-90 years (M = 74.9, SD = 8.0) completed a priming experiment exactly the same as that described above except that the prime duration was increased from 100 ms to 450 ms followed by a shorter 50-ms delay and then the target (see Table 14 for summary of changes). This was chosen so that the SOA was still 500 ms.

There was no evidence of priming as both integrative and semantic priming effects were negative with a high level of variability across participants (M = -19.04 ms, SD = 194.67, and M = -89.71 ms, SD = 270.88, respectively). A Wilcoxon Signed-ranks test was conducted to evaluate whether integrative or semantic reaction times were significantly different to baseline reaction times. For integrative primes, there was no significant difference from baseline, Z = -0.56, *ns*, and for semantic primes, there was no significant difference from baseline, Z = -0.84, *ns*.

As there was still no evidence of priming from these data, the salience of the prime was further increased in a second pilot study.

Table 14

Overview of the Differences Between Each Priming Experiment

Experiment	Fixation	Prime	Delay	Target	SOA
Experiment 5a	500 (red)	100 (red)	400	∞ (white)	500
Pilot 1	500 (red)	450 (red)	50	∞ (white)	500
Pilot 2	500 (red)	950 (white)	50	∞ (white)	1000
Experiment 5b	500 (red)	950 (red)	50	∞ (white)	1000

Note. Cells Show Duration in ms (Colour of Item).

Pilot 2. A further seven older participants (6 female), aged 66-86 years (M = 73.5, SD = 8.10), completed a similar priming experiment with the exact same procedure as above except that the prime was displayed in white for 950 ms and then followed by a delay of 50 ms, producing a 1000-ms SOA (see Table 14 for comparison). A white prime was used to increase the brightness of the prime against the black background. Originally eight participants were tested but one participant was excluded for failure to follow instructions.

This time there was a strong indication of priming for both integrative and semantic primes (M = 373.3 ms, SD = 385.2, and M = 397.2 ms, SD = 422.9, respectively). A Wilcoxon Signed-ranks test was conducted to evaluate whether integrative or semantic priming was significant. For integrative primes, there was a significant difference from baseline, Z = -2.36, $p < .05$, and for semantic primes there was a significant difference from baseline, Z = -2.36, $p < .05$. It is immediately obvious that this large amount of priming is unlikely to be entirely attributable to a priming effect. Participants completing the experiment experienced a lack of distinction between primes and targets because they were both white – the excluded participant was responding to primes as well as to targets. This was evidenced through increased errors during practice and a greater requirement to practice before commencing the main experiment. The main cause of this result however was due to the baseline condition – participants were often expecting two words to appear and to respond to the second (only 1/3 of trials were baseline primes). When the baseline asterisks appeared first, they often paused slightly on the following target word due to expecting a second word to appear. One participant specifically stated that she was making this mistake. This produced an 'asterisk effect' in the data.

The asterisk effect was measured by splitting the previously unanalysed nonword target data into nonwords preceded by asterisk primes and nonwords preceded by word primes. There was evidence of an asterisk effect as responses following asterisks primes were slower than those following word primes. The mean difference between the conditions was 181.96 ms (SD = 215.96). A Wilcoxon Signed-ranks test confirmed that the asterisk effect was significant, $Z = -2.03$, $p < .05$.

It was clear that to avoid the asterisk effect, a distinction needed to be made between targets and primes. This would prevent participants using the number of words they saw as a cue to respond. In the main part of Experiment 5b, the prime was returned to red but the SOA was kept at 1000 ms.

Method

Participants. Eighteen older adults (13 female), aged 61-85 years (*M* = 73.2, *SD* = 6.9), took part in the experiment.[26] They were recruited from the University of Warwick Age and Memory Study volunteer panel that was populated by local advertisements; they were offered no financial incentives for participation.

To assess cognitive functioning, participants completed the Digit Symbol Substitution task (Wechsler, 1981) as a measure of perceptual-motor or processing speed (*M* = 47.06, *SD* = 9.30). They also completed the multiple choice part of the Mill Hill vocabulary test (Raven et al., 1988) as a measure of crystallised intelligence (*M* = 23.72, *SD* = 3.61).

Materials and Procedure. The materials and procedure were identical to the Estes and Jones (2009) replication in Experiment 5a above except that the prime was displayed for 950 ms followed by a 50 ms delay (1000-ms SOA). There were six participants in each counterbalancing condition who saw different combinations of prime-target word pairs.

Results

As with Experiment 5a, only the correct responses for the lexical decision task were used to formulate response speed averages for each trial type and responses falling outside of 2.5 standard deviations from each prime type average were excluded as outliers.

[26] One 87-year-old participant was excluded from the analysis and replaced by another participant in the same counterbalancing condition. This is because he was more than two standard deviations away from the mean and sometimes more than three standard deviations from the mean on several different measures of performance.

To assess any effects of counterbalancing, a 3 (Prime type: integrative, semantic, baseline) x 3 (List type: 3 levels of counterbalancing between subjects) repeated measures ANOVA was conducted on the reaction time data to word targets for older adults. There was no main effect of list type, $F < 1$, and no interaction between prime type and list type, $F < 1$. This demonstrates that the counterbalancing did not influence the pattern of results. There was a main effect of prime type, $F(2, 30) = 25.62$, $MSE = 7{,}902$, $p < .001$, indicating priming effects because baseline reaction times were slower than integrative and semantic reaction times (see Figure 20).

Measures of integrative and semantic priming were produced by subtracting the mean reaction time for targets following integrative and semantic primes from reaction times for words following baseline primes (see the top of Figure 21). One sample *t*-tests were conducted to establish priming effects. There was a significant integrative priming effect, $t(17) = 5.35$, $p < .001$, and a significant semantic priming effect, $t(17) = 5.64$, $p < .001$, with targets following integrative and semantic primes showing faster reactions than targets following the baseline (asterisks prime) condition. In addition to this, a measure of any possible asterisk effect was also taken by subtracting reaction times for nonword targets following real word primes from nonword targets following asterisks primes (see Figure 20 for means). There was no significant asterisk effect, $t(17) = 1.36$, $p = .19$.

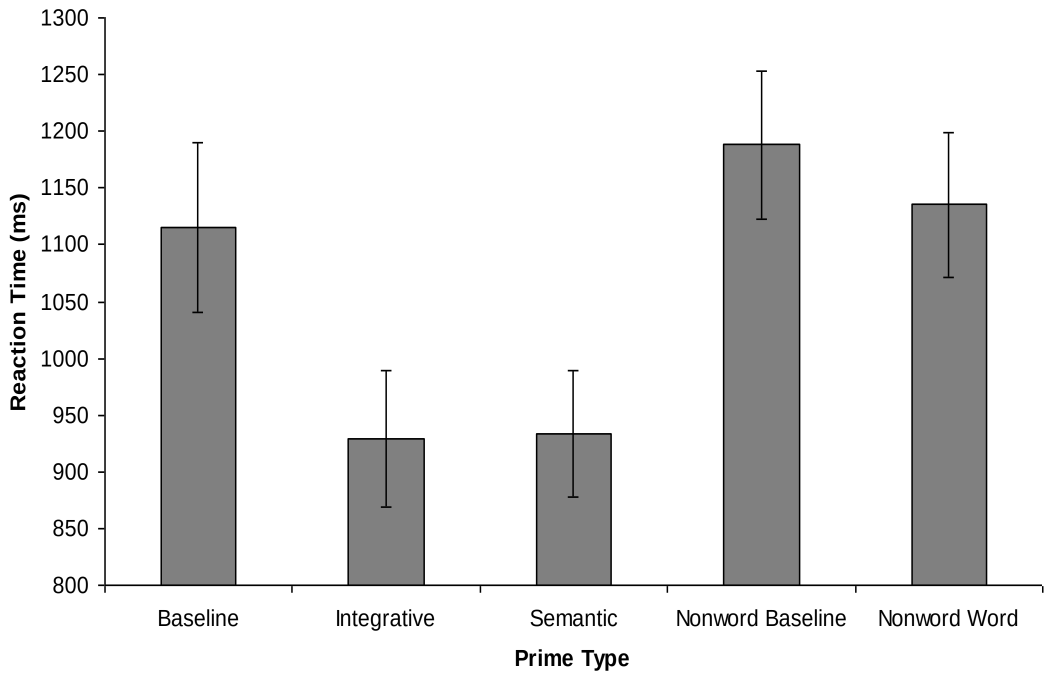

Figure 20. Mean correct reaction times to targets following different prime types: Baseline, integrative and semantic are primes preceding word targets. Nonword baseline and nonword word are baseline (asterisks) primes and word primes preceding nonword targets. Error bars are ± 1 *SE*
.

In order to fully establish that integrative and semantic priming effects were significantly above the asterisk effect, paired sample *t*-tests were conducted. The bottom of Figure 21 shows the integrative and semantic priming effects with the asterisk effect subtracted. Integrative priming effects were significantly larger than zero after subtracting the asterisk effect, $t(17) = 3.16$, $p < .001$. Semantic priming effects were also significantly larger than zero after subtracting the asterisk effect, $t(17) = 3.60$, $p < .01$. As both integrative and semantic priming were present in the data, a final paired sample *t*-test was conducted to establish if they were of different magnitude. There was no significant difference in the magnitude of the integrative and semantic priming effects, $t(17) = 0.25$, *ns*.

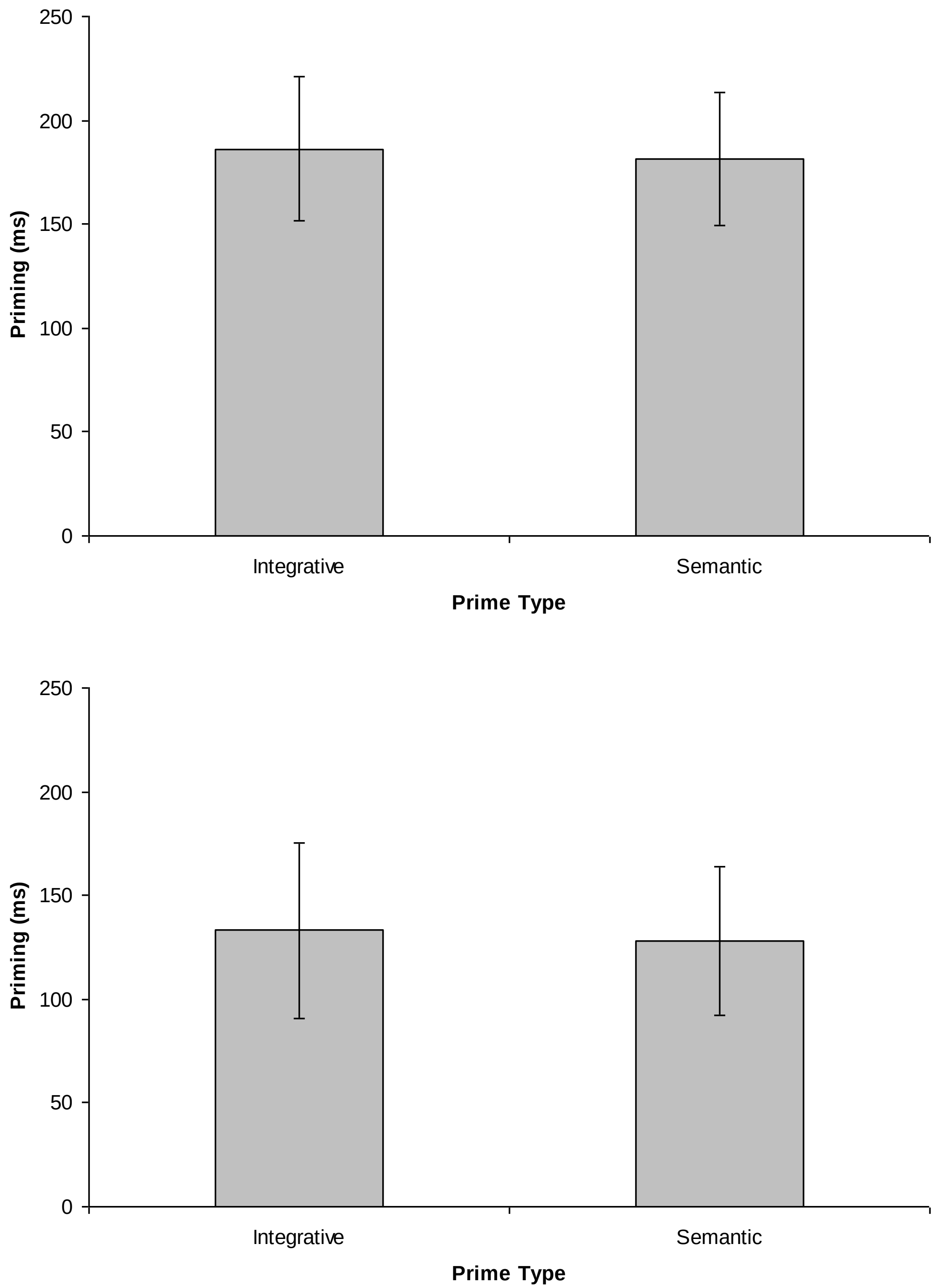

Figure 21. Integrative and semantic priming effects. Top: Integrative and semantic priming effects above baseline condition. Bottom: Integrative and semantic priming effects after subtracting the asterisk effect. Error bars are ± 1 *SE*.

As with Experiment 5a, the percentage of correct responses to word targets was also analysed for each prime type. The means were all identical and close to ceiling performance: Mean percentage correct was 99.6% for integrative, semantic and baseline responses.

Discussion

Experiment 5b demonstrated an integrative priming effect in older adults that was not significantly different to the size of the semantic priming effect. Estes and Jones (2009, Exp. 2) also found no difference in the overall magnitude of integrative priming compared to semantic priming with young adults. Although the current experiment did not have enough power to statistically differentiate between the semantic and integrative priming magnitudes, the presence of integrative priming was reliably established. This indicates that preexisting relations linking prime-target pairs (e.g., shared semantic features) are not necessary to elicit priming effects in older adults.

Integrative compounds form a vital part of language by reducing the number of words required to convey a specific concept. For example, a *plastic toy* is a more concise way of saying a “toy made from plastic”. Such compounds are common in language and they are useful for accelerating the communication of information. It is therefore perhaps unsurprising that such relationships facilitate the comprehension of a target word following an integrative prime. Language comprehension is largely unaffected by the ageing process (e.g., Burke, MacKay, & James, 2000) and it was noted earlier that semantic priming is present in older adults (Laver & Burke, 1993). The presence of integrative priming with older adults as well as young adults is therefore consistent with these observations.

Experiment 5b demonstrated that integrative priming occurs in older adults just as it does in young adults (Estes & Jones, 2009). Experiment 6 therefore tests whether integrative relations also facilitate memory in older adults just as they do in young adults (Jones et al., 2008). More specifically, Experiment 6 compares age differences in associative memory for integrative, semantic and unrelated word pairs. The experimental procedure was based closely on the cued recall element of Naveh-Benjamin's (2000) Experiment 4. If integrative relations alleviate the age-related deficit like semantic relations do (e.g., Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2005; Naveh-Benjamin, Hussain et al., 2003), this would indicate that stimuli that assist encoding and retrieval strategies can improve associative memory. Alternatively, if integrative relations fail to alleviate the age-related deficit, this would suggest that preexisting relations (shared features) between items are more important for supporting associative memory formation in older adults.

Experiment 6

Method

Participants. Thirty-six young adults (30 female) aged 18-32 years (*M* = 19.5, *SD* = 2.7) and 36 healthy older adults (20 female) aged 61-86 years (*M* = 73.1, *SD* = 6.9) took part in the experiment. Young participants were undergraduates at Warwick University (UK) who participated in exchange for course credit. Older participants were recruited from the University of Warwick Age and Memory Study volunteer panel that was populated by local advertisements; they were offered no financial incentives for participation. None of the participants had previously taken part in Experiment 5a or 5b.

To assess cognitive functioning, participants completed the Digit Symbol Substitution task (Wechsler, 1981) as a measure of processing speed. They also completed the multiple choice part of the Mill Hill vocabulary test (Raven et al., 1988) as a measure of crystallised intelligence. The results were consistent with the literature (e.g., Salthouse, 1991, 2010). Young participants were significantly faster at the Digit Symbol Substitution task, $t(70) = 11.02$, $p < .001$ (young $M = 72.9$, $SD = 10.2$; older $M = 45.5$, $SD = 10.9$). For the vocabulary test, young participants scored significantly lower than older participants, $t(70) = 5.89$, $p < .001$ (young $M = 15.6$, $SD = 4.2$; older $M = 21.9$, $SD = 4.8$).

Materials. The main memory stimuli were taken from a set of 180 words formed from four groups of 45 words (see Appendix 4). There were 45 target words, each paired with a corresponding integrative, semantic and unrelated cue word. This produced three sets of 45 cue-target pairs, all with the same 45 target words. For example the target word ‘book’ could appear in one of three combinations: integrative – *travel book*, semantic – *article book*, or unrelated – *lapel book*. In the experiment, participants would see two words; they would later be cued by being shown the left word of each pair and would be asked to recall the corresponding target word. Stimuli were arranged so that each participant would only see each and every target word once. Therefore every participant was recalling the exact same words, but not necessarily from the same cues (see details of counterbalancing below).

The target, integrative and semantic words were taken from Estes and Jones (2009) and were the same words as used in Experiment 5. The unrelated words were chosen such that they would be unrelated to their corresponding target words yet have similar length and frequency of occurrence in the English language as target, integrative and semantic words. Target, integrative and semantic words were grouped together and compared to unrelated words: Non-significant *t*-tests revealed that these two sets of words were of similar length, $t(177) = 1.20$, *ns*, and frequency of occurrence, $t(66.37) = 1.33$, *ns*, using log HAL frequency (Lund & Burgess, 1996). Twelve additional pairs of words, four of each category (integrative, semantic and unrelated) were created to be used as buffers and for a practice test.

Procedure. Stimuli were arranged into blocked sets each consisting entirely of integrative, semantic or unrelated pairs. Each block contained 15 pairs of words from the memory stimuli as well as two additional pairs (with the same type of relationship – integrative, semantic or unrelated), one at the start and one at the end, which were used as buffers. A total of 17 pairs were therefore displayed to participants for each memory test. Participants completed a separate memory test for each of the three pair types. Word pairs were presented in black with a white background in the centre of a laptop computer screen. Words were presented in lower case with a font size of 40 pt with a height corresponding to roughly 1.4º viewing angle at a distance of 60 cm.

Pilot studies were conducted with young adults to determine the optimal presentation of the memory set pairs. To avoid both ceiling effects for young adults and floor effects for older adults, the main experiment presented stimuli at a rate of 5 s per pair for young participants but 10 s per pair for older participants. These are the same presentation rates as used in Naveh-Benjamin's (2000) Experiment 4.

Before the main memory tests, participants completed a practice version of the experiment which presented six pairs of words sequentially (two of each relationship type, integrative, semantic and unrelated). Participants were informed that they would be required to memorise the words in each pair and that later they would be shown the left word of each pair and would be required to recall the corresponding right word. Practice pairs were shown sequentially at the same rate as the main experiment.

After the presentation of the last pair there was a 1-minute delay which was filled with counting backwards in threes from 200. Following this, a single cue word (which was always the left word of each pair) was shown on the screen. Participants were required to say the corresponding target word for each cue word and their responses were noted by the experimenter. After each response the next cue word was shown on the screen by the experimenter pressing a button. Cue words appeared in a randomised order for each participant.

In the main experimental procedure, the entire memory task was completed three times, once with each type of word pair relationship. In each case, participants viewed a sequential memory set of 17 pairs (15 pairs for the cued recall test and two buffers) at 5 s per pair for young participants and 10 s per pair for older participants. This was followed by a delay and then a cued recall test, which was conducted in the same way as described for the practice. Participants were offered the chance to rest between conditions.

Counterbalancing and randomisation was conducted throughout the experiment. Crucially, the condition order was fully counterbalanced so that every possible order of integrative, semantic and unrelated test was covered (six combinations of condition order). Furthermore, the target words were matched to different combinations of integrative, semantic and unrelated cue words in six different lists. This produced a 6 x 6 design such that no participants within each age group received the same conditions with the same stimuli in the same order. There were 36 different test combinations and one participant from each age group completed each one. Within experimental blocks, individual stimuli were presented in randomised order both during presentation and during cued recall.

Results

To begin with, to assess the effects of counterbalancing, a 2 (Age: young, older) x 3 (Condition: integrative, semantic and unrelated pairs) x 6 (Condition order: see above) repeated measures ANOVA was conducted on the cued recall performance data.[27] There was a marginal effect of condition order, $F(5, 60) = 2.20$, $MSE = 0.07$, $p = .07$, with lower performance when the semantically related pairs test was the second test to be completed. There was also an interaction between condition and condition order, $F(10, 120) = 2.01$, $MSE = 0.02$, $p < .05$, mainly because of a drop in integrative test performance in the condition order unrelated, semantic then integrative. There was also an interaction between condition order and age, $F(5, 60) = 2.69$, $MSE = 0.07$, $p < .05$, where only older participants were adversely affected on the integrative test for the specific condition order unrelated, semantic then integrative. No other interactions were significant. A similar analysis was conducted to assess the effects of stimuli order: a 2 (Age: young, older) x 3 (Condition: integrative, semantic and unrelated pairs) x 6 (Stimulus order: see above) repeated measures ANOVA was conducted on the cued recall performance data. There was no main effect of stimulus order, $F < 1$, and no interactions involving stimulus order. Overall the results here show that there were some

[27] There were insufficient residual degrees of freedom to conduct a 2 (Age) x 3 (Condition) x 6 (Condition order) x 6 (Stimulus order) repeated measures ANOVA.

different effects in performance due to condition order and this is considered the following analyses.

To assess whether integrative, semantic, and unrelated pairs were remembered differently between young and older participants, a 2 (Age: young, older) x 3 (Condition: integrative, semantic and unrelated pairs) repeated measures ANOVA was conducted on the cued recall data (see top of Figure 22). There was a main effect of age, $F(1, 70) = 27.95$, $MSE = 0.08$, $p < .001$, with older participants recalling significantly less than young participants. There was also a main effect of condition, $F(2, 140) = 147.71$, $MSE = 0.02$, $p < .001$, with performance in the unrelated condition being much lower than both the integrative and semantic conditions. The interaction between age and condition was also significant, $F(2, 140) = 13.86$, $MSE = 0.02$, $p < .001$. This is because although older participants performed lower than young participants in all conditions, the age difference was much larger for unrelated word pairs than for integrative and semantic word pairs. Despite performance levels being high for integrative and semantic pairs and low for unrelated pairs, the proportion of participants hitting ceiling and floor performance was low and comparable between young and older adults. For integrative, semantic and unrelated pairs the proportion of young adults performing at ceiling was .11, .17 and .03, respectively, and floor was 0 for all pair types; for older adults the proportion performing at ceiling was .17, .19 and .03, respectively, and floor was .03, 0 and .19, respectively. It is also important to note that although the integrative and semantic conditions yielded age differences that were very small, age differences were also reduced by increased presentation times for older adults. Therefore, integrative and semantic relations did not completely abolish age deficits; rather they reduced them relative to unrelated pairs.

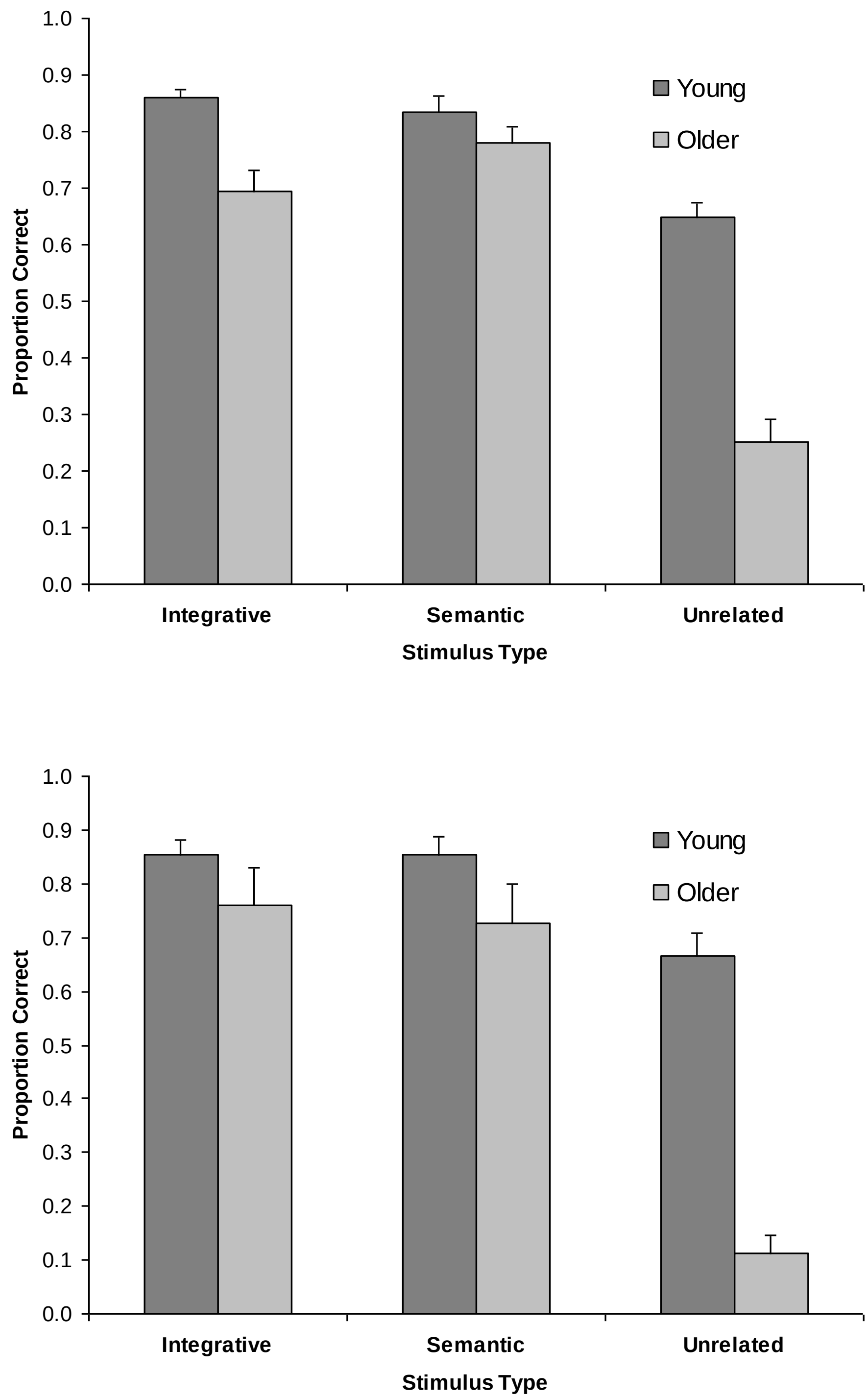

Figure 22. Young and older participants' performance for cued recall of integrative, semantic and unrelated word associations. Top: all data. Bottom: data from first test block only. Error bars are 1 *SE*.

Further tests revealed that there was no age by condition interaction between integrative and semantic conditions (*p* = .13), but the interaction was present between integrative and unrelated conditions (*p* = .001), and between semantic and unrelated conditions (*p* < .001). In order to establish if the lack of interaction between age and integrative and semantic memory performance was determinable, power analysis was conducted to measure the power present to detect this effect. The experiment had sufficient power to detect a medium size of effect for the interaction.[28] This means that if there is a difference in the effect of age between memory for integrative and semantic word pairs, it is likely to be only a small effect size.

In case of carry-over effects from one condition to another, the analysis was re-conducted using data only from the first condition that each participant completed (see bottom of Figure 22). Thus both age and condition were between subjects factors, with 12 young and 12 older participants in each condition. A 2 (Age: young, older) x 3 (Condition: integrative, semantic and unrelated pairs) factorial ANOVA revealed a qualitatively identical pattern of results. There was a main effect of age, $F(1, 66) = 42.59$, $MSE = 0.03$, $p < .001$. There was also a main effect of condition, $F(2, 66) = 47.66$, $MSE = 0.03$, $p < .001$, and an interaction between age and condition, $F(2, 66) =$

[28] The most informative estimate of power would not be based upon the effect size measured in the data, as it cannot be assumed to represent the effect size of the population as a whole (O'Keefe, 2007). As would be expected from the null result (O'Keefe, 2007), the power based upon the actual effect size measured was low: Power = .33. Power analysis was therefore conducted with G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007), using standard estimates of small and medium effect sizes taken from Murphy and Myors (1998). Power estimates were based on a repeated measures design and the correlation between integrative and semantic memory performance, $r(72) = .52$, $p < .001$, was used in the calculations. With 72 participants, and $\alpha = .05$, to detect a medium effect ($f^2 = .15$, $d = .5$) the experiment had a power of 1.00, and to detect a small effect ($f^2 = .02$, $d = .2$) the experiment had a power of .68. It is also worth noting that the main data had a larger age difference for integrative compared to semantic pairs but that the data from the first test only (see Figure 22, bottom) had the opposite pattern – larger age differences for semantic than integrative pairs, which further indicates no differential effect of stimuli type on memory performance across age.

13.97, *MSE* = 0.03, $p < .001$. This demonstrates that the overall pattern of results was not unduly influenced by a particular condition order.

Intrusions. Intrusions were categorised to ascertain if participants were aware of relationships between the word pairs they memorised. An intrusion was defined as a word response produced during the cued recall test that was not the correct answer. (Trials when participants made no response were categorised as omissions.) Intrusions were further coded on the basis of their congruence with the list type. For integrative and semantic lists, a congruent intrusion was when there was any relation between the cue and the intrusion. For unrelated lists, a congruent intrusion was when there was no relation between the cue and the intrusion. The classification of intrusions was conducted independently by two coders, both blind to the experimental condition and the age of participants. Initially the relatedness coding between coders was in agreement for 86% of intrusions – the remaining discrepancies were then resolved by discussion.

A 2 (Age: young, older) x 3 (Condition: integrative, semantic and unrelated test) x 2 (Congruency: congruent, incongruent) repeated measures ANOVA was conducted on the proportions of responses that were intrusions (see Figure 23 for means and Table 15 for summary of response types). There was a main effect of age, $F(1, 70) = 10.92$, *MSE* = 0.03, $p < .01$, with older participants producing more intrusions than young participants.[29] There was a main effect of condition, $F(1.72, 120.56) = 9.00$, *MSE* = 0.01, $p < .001$, with more intrusions for the unrelated condition than for the integrative or semantic condition. There was also a main effect of congruence, $F(1, 70) = 63.55$,

[29] Note that older adults produced around twice as many incorrect responses (intrusions plus omissions) as did young adults (see Table 15). Older adults also produced around twice as many intrusions as young adults; therefore the proportions of incorrect responses that were intrusions rather than omissions were approximately the same in the two age groups (*M* = 0.27, *SD* = 0.25, for young adults; *M* = 0.34, *SD* = 0.23, for older adults), $t(69) = 1.21$, $p = .23$.

MSE = 0.01, *p* < .001, with more congruent than incongruent intrusions. This is important as it shows that participants were aware of relations between words they were recalling. There was a significant interaction between age and congruency, *F*(1, 70) = 13.13, *MSE* = 0.01, *p* < .001. Both young and older participants made more congruent than incongruent intrusions but the difference was larger for older participants. There was also a marginal interaction between condition and congruency, *F*(1.38, 96.30) = 3.19, *MSE* = 0.02, *p* = .06. This was because there was a smaller difference between the number of congruent and incongruent intrusions for the unrelated test than for the integrative or semantic tests. Finally, the triple interaction between age, condition and congruency was not significant, *F*(1.38, 96.30) = 1.97, *MSE* = 0.02, *ns*.

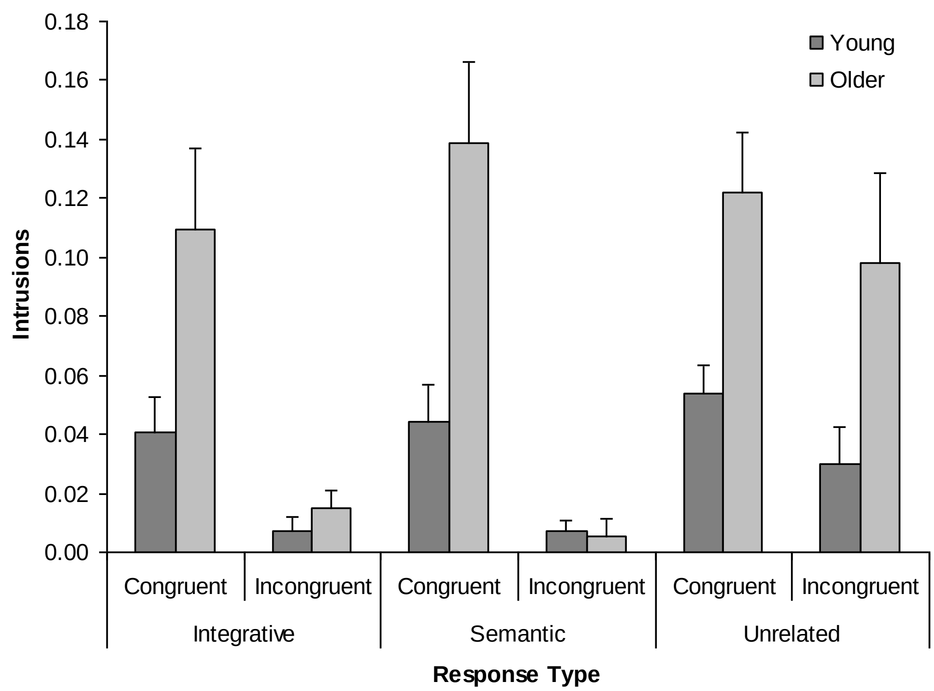

Figure 23. Mean proportion of responses that were intrusions, coded as congruent and incongruent with the test types for integrative, semantic and unrelated tests and for young and older participants. Error bars are 1 *SE*.

Table 15

Mean (and SD) Proportion of Correct, Intrusion and Omission Responses for Integrative, Semantic and Unrelated Conditions

		Response Type		
Condition	Age Group	Correct	Intrusion	Omission
Integrative	Young	.86 (.11)	.05 (.08)	.09 (.09)
	Older	.69 (.26)	.12 (.19)	.19 (.18)
Semantic	Young	.84 (.15)	.05 (.07)	.10 (.13)
	Older	.75 (.23)	.14 (.18)	.11 (.13)
Unrelated	Young	.59 (.20)	.08 (.12)	.33 (.18)
	Older	.24 (.24)	.22 (.22)	.54 (.26)

Preexisting relations. To examine the possibility that integrative word pairs had been encountered before and may therefore contain some preexisting relations, further analysis was conducted for each word pair within the integrative category: In total, the experiment used 45 different integrative word pairs. For each word pair, a measure of local co-occurrence was calculated using the British National Corpus (BNC, 2007) which is a collection of 100 million texts taken from written and spoken language. The database was used to calculate how frequently the individual words of each integrative pair occurred adjacently in the corpus of text. This measure of familiarity was highly suitable for integrative pairs as they are coherent when put together in language. Therefore, it provides an indication of the amount of prior exposure to links between the words. Across the 45 integrative word pairs there was a mean number of adjacent occurrences of 5.84 (*SD* = 11.44). That is, these word pairs occurred on average less than six times in 100 million texts.

In the experiment, each pair was tested with 12 young and 12 older participants, so for every pair there was a measure of both young and older participants' memory performance. The BNC co-occurrence measure was not significantly correlated with the proportion of correct responses for either young or older participants, $r(45) = -.01$, $p = .97$, $r(45) = .08$, $p = .58$, respectively. This indicates that for these items, amount of prior exposure did not affect memory performance. Within the integrative word pairs, there were 20 pairs that had no adjacent occurrences in the BNC, while the remaining 25 pairs had one or more occurrences. The memory data (proportion correct for each word pair) were therefore split and entered into a repeated measures ANOVA with co-occurrence as a 2-level independent factor (BNC co-occurrence: none, 1 or more) and age as a 2-level repeated factor (Age: young, older). Importantly there was no main effect of co-occurrence, $F < 1$, with no BNC co-occurrence memory performance ($M = .76$, $SD = 0.11$) showing similar levels to higher BNC co-occurrence memory performance ($M = .79$, $SD = 0.11$). There was a main effect of age, $F(1, 43) = 44.29$, $MSE = 0.01$, $p < .001$, with young participants ($M = .86$, $SD = 0.12$) recalling a higher proportion than older participants ($M = .69$, $SD = 0.15$). There was no interaction between BNC co-occurrence and age, $F < 1$. This indicates that memory performance for integrative word pairs was not attributable to preexisting relations in either young or older adults.

Discussion

The results replicated prior research: Older adults showed general memory deficits relative to young adults (e.g., Salthouse, 2010; Zacks et al., 2000) and more importantly the manipulation of semantic relatedness between to-be-associated words attenuated the age deficit in associative memory (e.g., Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2005). Unique to the current study was the manipulation of relatedness via integratability of words within word pairs. Similar to semantic relations (e.g., *paw foot*), integrative relations (e.g., *monkey foot*) alleviated the age-related associative memory deficit compared to that found for associations between unrelated words. Although integrative relations were known to facilitate memory among young adults (Jones et al., 2008), this is the first demonstration that such integrative relations also facilitate memory among older adults. Furthermore, integrative word pairs were similarly as powerful as the semantic word pairs in reducing the age-related memory deficit. Given that the words in integrative pairs were semantically dissimilar and unassociated, their attenuation of the age-related memory deficit cannot be directly attributed to preexisting relations. Rather they formed concepts that were consistent with world knowledge, which is perhaps why they could have been easier to encode and/or retrieve than unrelated word pairs. Analysis of intrusion errors also showed that participants were more likely to recall words with similar relatedness to the cue as the actual target word. For example, when prompted with a word from an integrative or semantic list, intrusions were more likely to be related to the cue. However, when prompted with a word from an unrelated list, intrusions were more likely to be unrelated to the cue. This is in line with previous research, where intrusions have been shown to share similar attributes to target stimuli (e.g., Underwood & Hughes, 1950). Also the use of a blocked design, which separated the three conditions (integrative, semantic and unrelated pairings), would have enhanced awareness of the relation types within each condition. Overall, the results from Experiment 6 provide strong evidence that integrative relations provide support for associative memory formation and/or retrieval, particularly among older adults.

General Discussion

Experiment 5b demonstrated that integrative priming was present in older adults. This established that integrative relations, like semantic relations, facilitate word processing among older adults. In Experiment 6, integrative, semantic and unrelated word pairs were used to assess cued recall performance in young and older adults. Age differences were significantly larger for unrelated word pairs than for both integrative and semantic word pairs. The reduction in associative deficits in older adults with integrative pairs therefore demonstrates a new type of support for associative memory performance in older adults. Previous research has suggested that semantic relations are easier for older adults to encode because fewer new connections need to be formed in memory (MacKay & Burke, 1990). This explanation cannot be applied to the integrative relations memorised in this experiment because the integrative word pairs were unassociated and semantically dissimilar. Instead, the results suggest that integrative word pairs may reduce associative deficits in older adults because they are easier to encode and perhaps more importantly easier to retrieve than unrelated word pairs. Furthermore, the guiding of encoding and retrieval processes could equally apply to semantically related word pairs where it is also easy to perceive relations between stimuli. Given that semantically and integratively related stimuli support associative memory performance in older adults to a similar extent and that integrative word pairs have no preexisting relations, the present study suggests that support for encoding and retrieval processes may be more important than preexisting relations for reducing age-related associative deficits.

Encoding. Integrative and semantic relations could alleviate the age-related memory deficit by inducing encoding strategies. Older adults are less likely than young adults to implement encoding strategies (e.g., Luszcz et al., 1990; Witte et al., 1990), and implementing encoding strategies has been shown to attenuate the age-related memory deficit (e.g., Naveh-Benjamin et al., 2007; Park et al., 1990; Treat & Reese, 1976). It is reasonable to conclude then that both integrative and semantic word pairs may show reduced age differences compared to unrelated word pairs because it is easier to meaningfully encode them. This conjecture is consistent with the popular view in cognitive ageing research that less effortful processes show smaller age-related decline (e.g., Fastenau, Denburg, & Abeles, 1996; Hasher & Zacks, 1979; Salthouse, 1988). It is also supported by the observation in Experiment 6 that participants' intrusion errors most often shared the same general relation to cues as the actual target items. Given that such occurrences were errors, the target words themselves clearly did not induce retrieval of the correct relation. Rather, it appears that the correct relation was retrieved but the correct item was not, thereby suggesting that the integrative and semantic relations might have been utilised as encoding strategies.

Retrieval. Alternatively, or additionally, integrative and semantic relations could alleviate the age-related memory deficit by inducing retrieval strategies. Indeed, there is evidence to suggest that associative deficits in older adults are a result of retrieval deficits more so than encoding deficits (Cohn, Emrich, & Moscovitch, 2008). In Naveh-Benjamin et al.'s (2005) Experiment 2, young and older adults completed an associative memory task with and without a secondary task to divide attention during recall. Young adults' recall performance was unaffected by dividing attention but older adults showed reduced memory performance with the presence of the secondary task. In contrast, Naveh-Benjamin et al.'s (2005) Experiment 1 showed that dividing attention during encoding affected both young and older adults equally. This evidence suggests that older adults may require more resources during recall. Naveh-Benjamin et al. (2005) also showed that performance on the secondary task dropped more for older adults than young adults during recall, especially when older adults were instructed to use memory strategies. This also indicates that older adults require more cognitive resources during associative memory recall. Finally, Naveh-Benjamin et al. (2007) found that encouraging participants to use encoding strategies reduced age-related associative deficits but encouraging participants to use encoding *and* retrieval strategies almost eliminated associative deficits in older adults.

The main demonstrations of associative deficits come from recognition tests of item and associative memory where older adults show smaller deficits for item than for associative memory compared to young adults (e.g., Naveh-Benjamin, 2000). The age-related deficits in associative recognition tests are often driven by increased false alarms to lures whilst endorsement of seen-before associations remains relatively intact (e.g., Castel & Craik, 2003; Healy et al., 2005). This means that older adults have formed associative memories but that they experience difficulty using recollection to reject lures. Therefore, this provides more evidence that encoding is intact in older adults and that it is retrieval that causes the age-related associative deficits observed.

The current results may thus be explained in terms of retrieval differences between the word pair types. The knowledge of relations between the words of integrative and semantic pairs during recall may have helped to narrow the search in memory for the corresponding target. It is well established in the literature that recognition tests yield smaller age differences than recall tests as they provide greater environmental support during retrieval (e.g., Craik & McDowd, 1987; Light et al., 2000; Naveh-Benjamin, 2000; Schonfield & Robertson, 1966). Therefore knowledge of the integrative and semantic relations during retrieval may have provided environmental support that benefited the older adults more than the young adults.

Integrative relations. In addition to demonstrating that the age-related memory deficit can be alleviated with previously unassociated word pairs, these experiments also contribute much to our understanding of integrative relations and their effects. Integrative priming has only recently been identified as a distinct influence on word processing (Estes & Jones, 2009), and similarly little research has examined the influence of integrative relations on memory (Jones et al., 2008). The present research demonstrates for the first time that integrative priming remains intact among older adults, and that integrative relations serve as powerful facilitators of memory across the lifespan. These findings are nontrivial, in that they contradict the view that older adults are disproportionately impaired at forming all types of new associations.

The introduction to this chapter discussed how, unlike semantic priming, integrative priming cannot be explained by pre-processing of the prime. Semantic priming may work by the prime activating features that are shared with the target, therefore allowing the target to be perceived faster/more easily. The same cannot be said for integrative relations because the relation is only apparent *after* the target appears (e.g., you would not naturally think of the target *foot* when presented with the prime *monkey*). This distinction provides a time course for how integrative relations may facilitate memory encoding in both young and older adults. The benefit of associative memory formation between integratively related words compared to unrelated words must occur after the relation has been perceived, indicating that it is only once a relation is established that it can aid memory formation. It may be the case that when encoding unrelated word pairs, age differences are large because older adults are less able to construct a structural context upon which to bind the words together. The key point is that it may not be the process of perceiving/generating a relation that aids memory, rather it is the use of relations *once perceived* to meaningfully encode stimuli. Therefore the age-related deficit in associative memory may lie in a deficit in generating meaningful relations between stimuli, not making use of those relations once generated.

Chapter 7: Does Strategic Support Reduce Associative Memory Deficits in Children Relative to Young Adults?

There is much research that shows cognitive development in children mirroring cognitive decline in older adults (e.g., Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Craik & Bialystok, 2006). This can apply to cognitive measures such as speed (Kail & Salthouse, 1994), executive functioning and conscious contributions to memory (Zelazo, Craik, & Booth, 2004), where children and older adults show deficits relative to young adults. Physiologically, young adults also have increased levels of white and grey matter volume compared to children and older adults (see Craik & Bialystok, 2006). This pattern of lifespan development and decline is particularly apparent in terms of episodic memory, which is the focus of this chapter.

Children are unable to form episodic memories in early childhood, resulting in childhood amnesia, where individuals cannot usually remember events they experienced before the age of around 4 or 5 years (e.g., Perner & Ruffman, 1995; Shing et al., 2010). Direct measures of episodic memory show young children to have deficits in reality monitoring (judging whether an action was performed or imagined) compared to older children (Lindsay, Johnson, & Kwon, 1991; Sluzenski, Newcombe, & Ottinger, 2004). This pattern is mirrored in older adults, who reliably show deficits relative to young adults in episodic memory measures (see Spencer & Raz, 1995, for a review). Episodic memory has also been shown to decline further throughout old age (Singer, Verhaeghen, Ghisletta, Lindenberger, & Baltes, 2003). It has been hypothesised that cognitive abilities that are last to develop are more complex than cognitive abilities that develop early and this causes them to be more susceptible to the ageing process (see Craik & Bialystok, 2006).

It has been shown earlier (Chapter 2) that episodic memory is closely related to associative memory, specifically, because the formation of episodic memory involves associations between units in memory (Trinkler et al., 2006). There is a large body of research into age-related associative deficits in older adults (Old & Naveh-Benjamin, 2008a), which correspond to age deficits observed in episodic memory. The nature of associative memory has been explored in much more detail in relation to ageing than development. There are, however, several studies that show associative deficits in children relative to young adults. This is in line with episodic memory deficits observed in children.

Cowan, Naveh-Benjamin, Kilb, and Saults (2006) examined associative deficits in children and older adults compared to young adults for associations between an object (item) and its spatial location. In Experiment 2a they found associative deficits in children and older adults compared to young adults: This is because associative memory deficits were larger than item memory deficits. However, in their Experiment 1a, children showed equivalent deficits for item and associative memory (older adults still showed increased associative deficits compared to item deficits relative to young adults). Sluzenski, Newcombe, and Kovacs (2006) also found associative deficits in young and older children relative to young adults. Participants were required to remember animals and the backgrounds in front of which they were presented. Relative to young adults, children showed a larger deficit in memory for the association between an animal and a background than memory deficits for the animals and backgrounds individually. Associative memory improved from young to older children. This is in line with a study by Lloyd, Doydum, and Newcombe (2009), who measured associative memory for objects and their backgrounds with young and older children. Young children showed similar performance to older children for objects but not for associations between objects and backgrounds. This only occurred in one condition based on long-term memory but a second working memory condition showed similar item and associative memory performance in both groups of children. Shing et al. (2008) measured memory for word pairs in children, young adults and older adults. Children and older adults had poorer memory for associations between words than young adults. However, no comparison was made to establish if associative memory deficits were different to item memory deficits. A similar study from the same laboratory showed the same result (Shing, Werkle-Bergner, Li, & Lindenberger, 2009). Finally, Lindsay et al. (1991) conducted a range of source monitoring experiments on children and young adults (e.g., remembering if a word was presented from a source on their left or their right). Children showed greater source memory deficits than young adults. Overall, the literature demonstrates that children are poorer at forming associations than are young adults, although there is limited evidence to establish whether this deficit is greater than memory deficits for items or units of information. In the current study, children and young adults were tested on both item and associative memory measures to establish how performance on the different memory tasks changes during development.

Previous lifespan studies have hypothesised that poor associative memory performance observed in children is due to strategic deficits rather than associative deficits whereas older adults' associative memory is affected by both strategic and associative deficits (Shing et al., 2010; Shing et al., 2008). Shing et al. (2010) proposed that strategic deficits in children occur because of protracted development of the prefrontal cortex. This observation coincides with Zelazo et al. (2004) who observed deficits in executive functioning in children relative to young adults. A neuroimaging review by Casey, Tottenham, Liston, and Durston (2005) also concluded that basic sensory functions develop first in children followed by more complex top-down cognitive functions. They found protracted development of prefrontal cortex (which they labelled as 'association cortex') relative to sensorimotor cortex. The current study therefore aimed to investigate the strategic component of associative memory formation by manipulating the semantic relatedness of to-be-associated stimuli. In the previous chapter, a reduction in associative deficits in older adults was obtained in cued recall by using semantically related word pairs. It is well established in the literature that semantic relations between to-be-associated stimuli reduce age-related associative deficits (e.g., Naveh-Benjamin, 2000, Exp. 4; Naveh-Benjamin, Hussain et al., 2003, Exp. 2). Given that this effect is reliable in older adults, the current study aimed to replicate this in children. In the previous chapter, the results indicated that semantic relations between to-be-associated stimuli may provide extra guidance at encoding and retrieval. Therefore, this may provide strategic support to participants which would establish if children differentially benefited from strategic guidance compared to young adults. Results from Shing et al. (2008) demonstrated that children's associative memory benefited more than young adults' associative memory by experimental manipulation of strategy use and the current study aimed to extend this finding.

In brief, the current study aimed to address two questions: 1. Do children show an associative memory deficit relative to young adults above and beyond deficits in item memory? 2. Is this deficit reduced by semantic relations between to-be-associated stimuli as is found in older adults? Both of these questions would help establish if associative deficits in children mirror those found in older adults, which would clarify the observed symmetry in the literature between episodic memory development and decline across the lifespan.

Experiment 7

Method

Design. The study was designed to test item and associative memory for semantically related and unrelated stimuli. The study used a similar method to Naveh-Benjamin et al.'s (2003) Experiment 2 but with different age groups of participants. Children and young adults studied word pairs and were tested on their recognition of the words (item memory) and the association between pairs of words (associative memory). Half of the word pairs comprised words that were semantically related to each other and the other half comprised unrelated words. Age was a between-participants factor (young children, older children and young adults) and memory test (item/associative) and relatedness of word pairs (related/unrelated) were within-participants factors.

Participants. The study contained three age groups and in total, 97 participants were tested.[30] There were 33 young children (15 female), aged 6-7 years (M = 6.94, SD = 0.24), 34 older children (17 female), aged 10-11 years (M = 10.88, SD = 0.33), and 30 young adults (15 female), aged 19-21 years (M = 20.40, SD = 0.68). The young and

[30] The data from this study were gathered by Charlotte Gallimore and Jennifer Harber.

older children were sampled from a primary school in a middle-class affluent area. Consent was gained from the head teacher and parents were notified. Parents were able to opt out their child from the study if they did not want the child to take part. The young adults were an opportunity sample of (non-psychology) students from the University of Warwick.

Materials. All of the words were chosen so that they would be appropriate to use with children by using the University of Essex Children's Printed Word Database (Stuart, Masterson, Dixon, & Quinlan, 1993-1996). This is a database of word frequencies in books available to children aged 5-9 years. Only words with a frequency greater than 50 occurrences per million were used in the study. The study phase of the memory experiment consisted of 42 word pairs, 21 pairs containing words that were semantically related and 21 containing words that were not semantically related. All of the words were taken from one of seven themed categories (see Appendix 5). Each of the seven categories produced three pairs of words. Related word pairs were constructed by taking two words from the same category and unrelated word pairs were constructed by taking two words from different categories. Relatedness was measured using a forward free association norms database (Nelson et al., 2004). For related word pairs, only associations with a value of 0.2 or above were used – this corresponds to a probability of 0.2 of generating the right-hand word of a pair (e.g., dog) when given the left-hand word (e.g., cat). For unrelated word pairs, there was no generation of the right-hand word at all when given the left-hand word in the database. In addition, an independent measure of relatedness was also obtained. Eleven undergraduate students completed an online questionnaire that contained a list of the related and unrelated word pairs. For each pair they rated their relatedness on a scale of 1 (unrelated) to 10 (closely related). The related pairs used in the study had a mean rating of 7 or above and the unrelated pairs had a mean rating of 4 or below.

The item recognition test consisted of 56 words: Half of the words were seen before in the memory set and half were completely new words (see Appendix 5). In addition, half of the seen-before words were originally in related word pairs and half were originally in unrelated word pairs (14 of each type). The associative recognition test consisted of 28 word pairs: Half were presented exactly as they were in the study phase (intact) and half were recombined, taking one word from one pair and the other word from another pair. Similar to the item test, half of the intact pairs were related pairs at study and half were unrelated pairs at study (seven of each type). Also half of the recombined pairs were related, using words from different related study pairs, but from pairs that were in the same related category (e.g., animals: study pairs *cat-dog* and *lion-tiger* recombined to form *cat-tiger*). The other half of the recombined words were unrelated and were all constructed from pairs that were unrelated at study. In the associative tests, words that were originally studied on the left of a pair remained on the left (and the same for words on the right). This ensured that memory was tested for the associations and not the spatial position of the words.

In total, three different sets of item and associative test lists were created using different combinations of old/new words following the same rules outlined above. This produced three different test lists for counterbalancing, which were randomly assigned to participants (three test list groups). The individual stimuli were balanced across the three test lists. For example, if the related pair *summer-spring* appeared at study, one list would contain *summer* in the item test, another list would contain *summer-spring* (intact) in an associative test, and the third list would contain *summer-winter* (recombined) in an associative test. Also, a given old item was not repeated across the two tests, that is, it appeared in either the item test or the associative test. Finally, the order of tests was counterbalanced so that half of the participants received the item test before the associative test and half vice-versa.

Words were presented in the centre of a laptop computer screen in a 40 point font corresponding to a height of approximately 1° at a 60-cm viewing distance. Word pairs at study and test were presented in lowercase with a dash between the two words. All words were displayed in a black font with a white background. Order of presentation of words (item test) and word pairs (study list and associative test) was randomised for each participant.

Procedure. All participants completed a short practice version of the memory test containing a study phase of three pairs and then item and associative tests. Participants were therefore explicitly aware of the experimental procedure before they began the main study. The 42 study pairs were presented sequentially at a rate of 5 s per pair. Following this, there was a distracter delay period for 60 s. The three groups of participants completed different tasks during this delay to prevent rehearsal. The tasks reflected their different abilities: The young children completed a two-times table, the older children counted backwards in threes from 100, and the young adults counted backwards in sevens from 200. After the delay, participants completed the item and associative memory tests. Test stimuli were shown on the screen and participants responded yes or no as to whether they had seen the test stimuli before. The test was self paced as only after each response did the next test stimuli appear. Young and older children responded verbally and the experimenter input the response. The young adults pressed the left (new/not-seen-before) and right (old/seen-before) mouse buttons themselves to advance. There was no gap between the item and associative tests apart from a screen to indicate the next test would begin. Throughout the session the experimenter read the words aloud to the youngest group of children only. This was to ensure that they were not hindered by their reading ability.

Results

Accuracy. Overall performance on the recognition memory tasks was ascertained by calculating hit rates minus false alarm rates. Proportions of hits and false alarms separately are reported in Table 16. [Note that for the item test, false alarm rates were based on new items that were neither related nor unrelated at study. Therefore, for the item tests, the same false alarm rate applies to the two different relatedness conditions.] Hits minus false alarms were calculated separately for the item and associative memory tests, and for test words that were related and unrelated at study. The signal detection measures d’ and ln(β) were also calculated as described in Chapter 5.

First, to assess any effects due to counterbalancing, a 3 (Age group: young children, older children, young adults) x 2 (Pair relatedness at study: related, unrelated) x 2 (Test type: item, associative) ANOVA was conducted on the hits minus false alarms data. Counterbalancing was assessed by adding both test list group as a 3-level between-subjects factor (test list groups 1, 2 and 3) and test order (item then associative or associative then item) as a 2-level between-subjects factor.

There was no main effect of test list group, $F(2, 79) = 1.41$, $MSE = 0.13$, *ns*. However, there was an interaction between pair relatedness at study and test list group, $F(2, 79) = 3.47$, $MSE = 0.04$, $p < .05$, with the difference between related and unrelated performance smaller for test list group 3 than for groups 1 and 2. There was also an interaction between test type and test list group, $F(2, 79) = 4.59$, $MSE = 0.05$, $p < .05$, with a larger difference in performance between the item and associative tests for test list group 1 than for groups 2 and 3. There was no main effect of test order, $F < 1$. There was a marginal interaction between test type and test order, $F(1, 79) = 3.96$, $MSE = 0.03$, $p = .050$, with performance higher for the item test when the item test was taken before the associative test and performance higher on the associative test when the associative test was taken before the item test. This may be a result of recency effects and/or interference in the second test caused by exposure to the first test. There were no other interactions involving test order. As the counterbalancing factors did not interact with age group, they were not considered in further analyses.

Table 16

Means and Standard Deviations for the Proportion of Hits and False Alarms

Test and group	Unrelated				Related			
	Hits		FA		Hits		FA	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Item								
Young children	.62	.20	.08*	.16*	.73	.18	.08*	.16*
Older children	.61	.20	.09*	.11*	.70	.18	.09*	.11*
Young adults	.73	.14	.10*	.10*	.80	.12	.10*	.10*
Associative								
Young children	.52	.30	.35	.30	.83	.20	.34	.27
Older children	.54	.24	.23	.19	.76	.20	.20	.20
Young adults	.61	.26	.25	.19	.74	.18	.20	.19

*FA based on new items is the same data for unrelated and related item tests.

A 3 (Age group: young children, older children, young adults) x 2 (Pair relatedness at study: related, unrelated) x 2 (Test type: item, associative) ANOVA was conducted on the hit minus false alarms data (see Figure 24 for means). There was a marginal effect of age, $F(2, 94) = 2.43$, $MSE = 0.12$, $p = .094$, with young children performing most poorly ($M = .47$, $SD = 0.22$) followed by older children ($M = .50$, $SD = 0.26$) and then young adults ($M = .56$, $SD = 0.25$). There was a main effect of pair relatedness at study, $F(1, 94) = 66.49$, $MSE = 0.04$, $p < .001$, with performance for words that were unrelated at study ($M = .43$, $SD = 0.26$) lower than performance for words that were related at study ($M = .59$, $SD = 0.23$). There was also a main effect of test type, $F(1, 94) = 85.96$, $MSE = 0.05$, $p < .001$, with performance higher for item memory ($M = .61$, $SD = 0.19$) than for associative memory ($M = .41$, $SD = 0.30$). There was an interaction between test type and age group, $F(2, 94) = 3.28$, $MSE = 0.05$, $p < .05$, with young children showing a larger difference between item and associative memory ($M = .27$) than older children ($M = .13$) and young adults ($M = .22$). Therefore, there was some evidence of an associative deficit in young children relative to older children and young adults. Finally there was an interaction between pair relatedness at study and test type, $F(1, 94) = 17.57$, $MSE = 0.04$, $p < .001$, with associative memory benefiting more from relatedness ($M = .25$) than item memory ($M = .09$). There were no other interactions in the data. With d' the results were qualitatively identical except that the main effect of age was no longer marginal, $F(2, 94) = 1.02$, $MSE = 1.40$, *ns*.

There was no three-way interaction between age group, pair relatedness at study, and test type. Therefore the difference between item and associative memory did not reduce more for children than for young adults for words that were in semantically related word pairs at study compared to words that were in unrelated word pairs at study. That is, associative memory deficits relative to item memory deficits in children were not alleviated by semantic relations between to-be-associated stimuli. However, planned comparisons showed that some trends were as expected: For unrelated word pairs, associative memory in young children was significantly worse than older children,[31] $t(65) = 1.98$, $p = .05$, and young adults,[32] $t(61) = 2.33$, $p < .05$, whereas for semantically related word pairs at study, there were no age differences in associative memory, $F < 1$. Again, with d' the results were qualitatively identical.

[31] Levene’s test for equality of variances was non significant so equal variances were assumed.
[32] Levene’s test for equality of variances was significant so equal variances were not assumed.

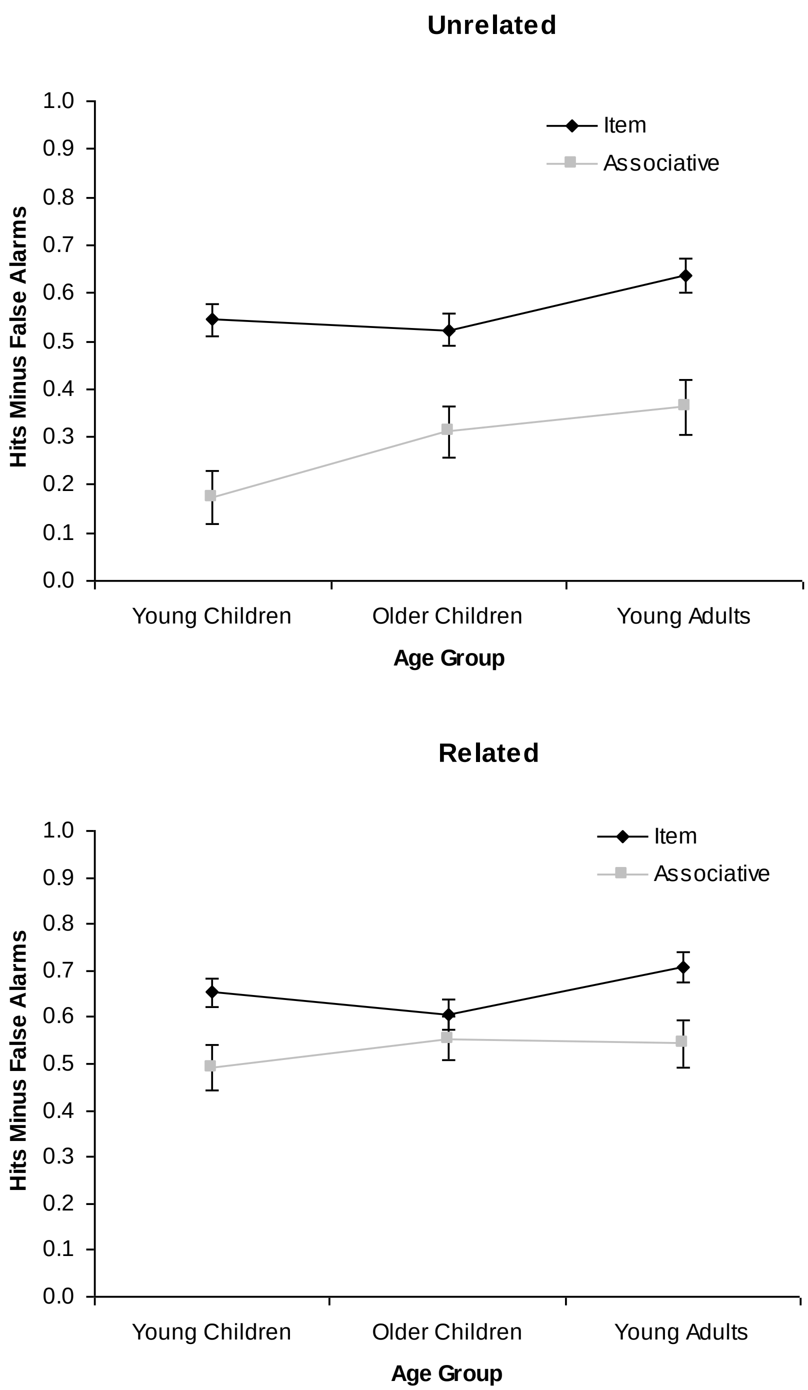

Figure 24. Hit minus false alarm rates for item and associative memory recognition tests for young children, older children and young adults for words that were in unrelated (top) and related (bottom) word pairs at study. Error bars are ± 1 *SE*.

Response Bias. A 3 (Age group: young children, older children, young adults) x 2 (Pair relatedness at study: related, unrelated) x 2 (Test type: item, associative) ANOVA was conducted on the response bias ln(β) data. Table 17 shows the means: Note that positive values of ln(β) represent a bias towards responding no/not-seen-before and negative values represent a bias towards yes/seen-before in the recognition tests. There was no main effect of age, $F(2, 94) = 1.32$, $MSE = 1.09$, *ns*. There was a main effect of pair relatedness at study, $F(1, 94) = 9.11$, $MSE = 0.19$, $p < .01$, with a stronger bias towards no/not-seen-before for pairs that were unrelated at study (M = 0.61, SD = 0.69) than for pairs that were related at study (M = 0.48, SD = 0.78). There was also a main effect of test type, $F(1, 94) = 107.08$, $MSE = 0.80$, $p < .001$, with a stronger bias towards no/not-seen-before for the item test (M = 1.01, SD = 0.94) than for the associative test (M = 0.07, SD = 0.53). There was also an interaction between age group and test type, $F(2, 94) = 4.68$, $MSE = 0.80$, $p < .05$, with a larger difference in response bias between item and associative tests for young children (M = 1.31) than for older children (M = 0.89) and young adults (M = 0.63). There were no other interactions in the data.

Table 17

Response Bias (ln(β)) for Item and Associative Tests for Unrelated and Related Pairs at Study for the Three Age Groups

Age Group	ln(β)							
	Unrelated				Related			
	Item		Associative		Item		Associative	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Young Children	1.34***	0.97	0.09	0.50	1.16***	1.05	-0.22*	0.54
Older Children	1.11***	0.95	0.23**	0.46	1.01***	0.92	0.11	0.58
Young Adults	0.82***	0.74	0.07	0.44	0.64***	0.87	0.14	0.62

Note. Significance indicates differences from zero.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Reaction times. Mean reaction times (RTs) were calculated for the different response categories: hit, miss, false alarm and correct rejection (see Figure 25). RTs were the time taken to respond to a test trial following onset of the test stimulus. RTs greater than 10 s were excluded from the analysis (0.4% of the data). Note that children responded verbally to test stimuli (then the experimenter entered their response) and young adults responded via a button press. Therefore, some age differences were expected because for children (but not young adults) there would be an extra delay between the experimenter hearing their response and entering it into the computer. It can be seen from the figure that RTs were generally slowest for young children then older children and then young adults. It is also clear that RTs to pairs that were unrelated at study were slower than RTs to pairs that were related at study. RTs to item tests were generally faster than RTs to associative tests. Additionally, correct responses were generally faster than incorrect responses for both endorsements (hits vs. false alarms) and rejections (correct rejections vs. misses). This indicated that there was no speed accuracy trade off because accuracy was higher for the faster responses. Not every participant made a response in every response category so this pattern of results was confirmed by analysing hit RTs only below.

A 3 (Age group; young children, older children, young adults) x 2 (Pair relatedness at study; related, unrelated) x 2 (Test type; item, associative) ANOVA was conducted on the hit RTs only (see Figure 25 for means). There was a main effect of age, $F(2, 90) = 57.88$, $MSE = 8.37 \times 10^5$, $p < .001$, with young children responding the slowest (M (SD) = 2846 (762) ms), then older children (M (SD) = 2103 (531) ms) and then young adults (M(SD) = 1572 (611) ms). Even though children and young adults responded differently, the difference is larger than can be accounted for by the experimenter's reactions when inputting the children's responses. Also the difference between young and older children of 743 ms is genuine as they both responded in the same way. There was a main effect of pair relatedness at study, $F(1, 90) = 10.47$, $MSE = 1.94 \times 10^5$, $p < .01$, with RTs to unrelated pairs at study (M (SD)= 2248 (858) ms) slower than RTs to related pairs at study (M (SD)= 2099 (831) ms). There was also a main effect of test type, $F(1, 90) = 76.68$, $MSE = 3.54 \times 10^5$, $p < .001$, with RTs to item test hits (M (SD)= 1903 (730) ms) faster than RTs to associative test hits (M (SD)= 2444 (959) ms). There was also an interaction between test type and age group, $F(1, 90) = 3.68$, $MSE = 3.54 \times 10^5$, $p < .05$, with a larger difference between item and associative RTs for young children (M = 735 ms) than older children (M = 333 ms) and young adults (M = 558 ms).

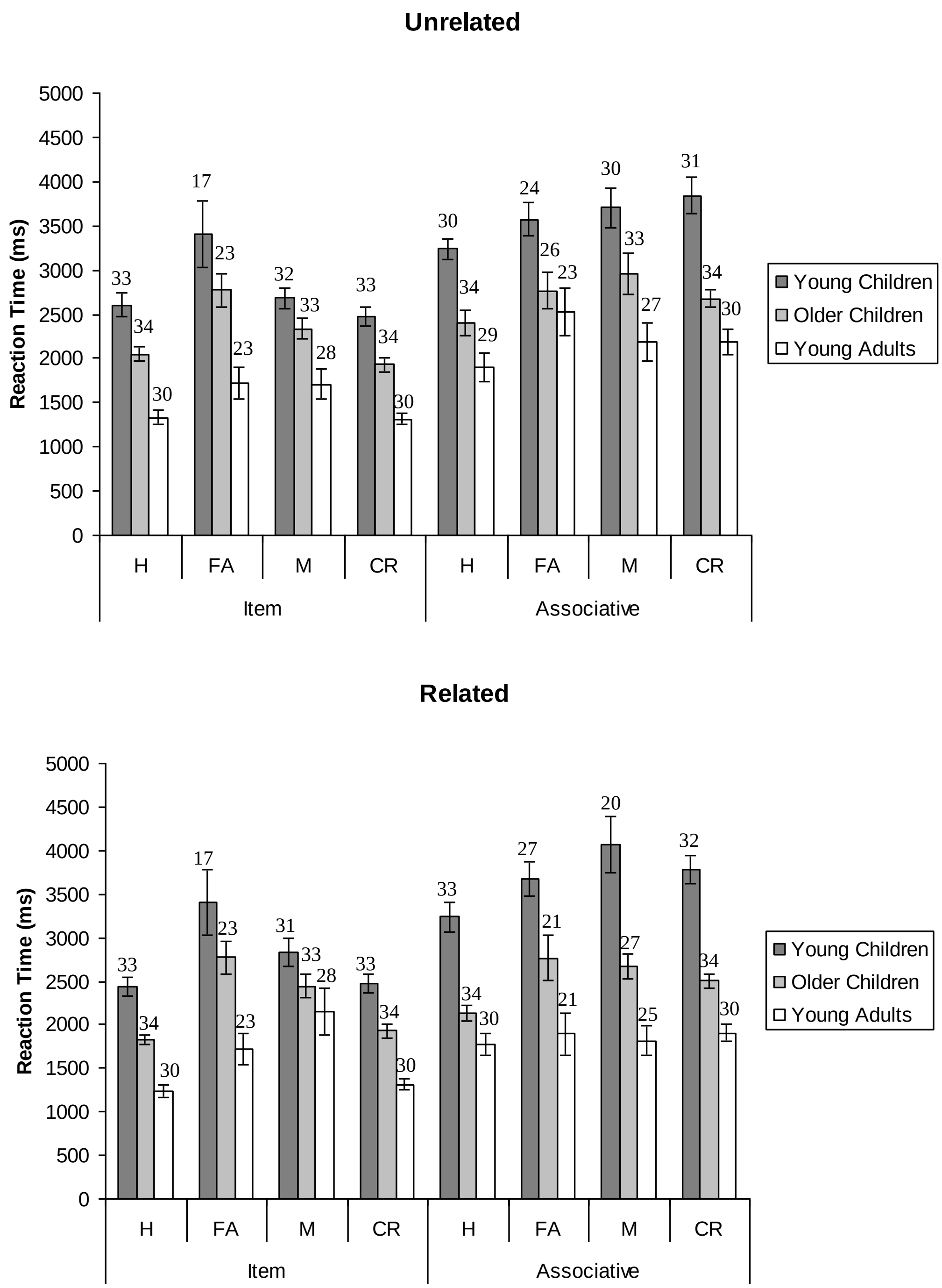

Figure 25. Mean reaction times for each recognition response type: hit (H), false alarm (FA), miss (M), and correct rejection (CR). Data shown for item and associative recognition in young children, older children and young adults for words that were unrelated (top) and related (bottom) at study. Error bars are ± 1 *SE*. Numbers show the number of participants contributing to each response type. Note that data for false alarms and correct rejections are the same for related and unrelated item tests.

Discussion

In general, children showed poorer memory than young adults, which is consistent with previous research. Contrary to expectations, associative memory relative to item memory in children was not statistically worse than young adults. However, there was a numerical trend in this direction, indicating associative deficits in young children relative to young adults. Furthermore, RTs also pointed towards associative deficits in children – for hit responses (correct recognition endorsements) the difference in RTs between item and associative tests was larger for young children than young adults. Therefore, associative tests required more consideration before responding for young children than young adults. One of the clearest effects in the data was the effect of relatedness. When word pairs were related at study and test, associative memory was enhanced in all age groups. There was also an effect of relatedness on RTs, with pairs related at study generally resulting in faster responses than pairs that were unrelated at study.

The presence of greater associative than item memory deficits in children relative to young adults in previous research is less robust when re-considered. The current study showed that if present, the effect is small and not comparable to the age-related associative deficit observed when young and older adults are compared. Cowan et al.'s (2006) Experiment 1 did find this effect in children but it was not present in Experiment 2. The associative memory measure in their study was unusual: both the item memory measure and the associative memory measure involved recognising the change in colour of an item. Therefore the item and associative measures were very similar. Sluzenski et al. (2006) also found associative memory deficits in children. A difference between their study and the current study was that they used very young children (4 and 6 year olds). They also found an improvement in relational (associative) memory between the 4 and 6 year olds as did Lloyd et al. (2009). It may therefore be the case that the children used in the current study were old enough to have a developed episodic and associative memory capability. It was shown in the introduction that childhood amnesia occurs mainly before the age of 6 years (e.g., Perner & Ruffman, 1995). Therefore, greater associative than item memory deficits in children relative to young adults may only occur with very young children. This poses a problem for future research because very young children may have difficulty following instructions and performing above floor if completing the same task as young adults. More importantly, if associative memory ability forms at a very young age, this indicates an asymmetry between associative memory development and decline because associative memory ability declines steadily throughout old age, not just at very old age (Bender, Naveh-Benjamin, & Raz, 2010).

The manipulation of relatedness did not seem to differentially affect children compared to young adults. The results are therefore not in line with hypotheses that suggest associative memory in children is reduced by deficits in strategy utilisation (Shing et al., 2010; Shing et al., 2008). This is because relations between words are likely to support strategy use by helping participants to form meaningful relations between the words of each pair at study and there was little evidence that this was occurring more for children than for young adults. The application of strategy and executive thought are commonly associated with the prefrontal cortex (e.g., Zelazo et al., 2004) which is yet to be fully developed in young children (e.g., Casey et al., 2005;

Craik & Bialystok, 2006). Therefore, the greater benefit of strategy support in children compared to young adults is consistent with the notion of frontal deficits in childhood. Also, M. K. Johnson, Hashtroudi, and Lindsay (1993) argue that frontal development is responsible for source memory deficits in children, which suggests links between frontal areas and episodic/associative memory formation. This view is consistent with the ageing literature, where age-related episodic deficits are linked to frontal decline (e.g., Mitchell, Johnson, Raye, Mather et al., 2000; West, 1996). This neuropsychological data can also be considered alongside the notion that associative deficits in children only occur at a very young age. M. H. Johnson (2001) notes rapid development of the brain in very young children with the overall structure not resembling that of adults until 3 years of age. This also indicates that very young children may have greater susceptibility to associative deficits than the children who participated in the current study. Very young children may therefore benefit more from relatedness than older children.

There was some evidence that the young children in the current study performed slightly better than the older children as can be seen in the item tests in Figure 24. Many studies have shown similar results where memory in children at around 10-13 years old shows a *developmental dip* relative to young children and young adults. M.S. Chung and Thomson (1995) reviewed the developmental dip in relation to memory for faces. For example, Carey, Diamond, and Woods (1980) found that memory for unfamiliar faces improved from ages 6 to 10 but then remained stable or declined before improving again at age 16. They argued that hormonal upheavals at puberty may disrupt memory performance. Flin (1980) also found a improvement in unfamiliar face identification from 6 to 10 years but a deterioration at 11 and 12 years with performance regained at

13 years. The effect is also found with other stimuli: Flin (1985) observed a slight developmental dip in flag recognition at age 13 years and also a dip in house recognition from 12-13 years. Similarly, Mann, Diamond, and Carey (1979) found a developmental dip in voice recognition, and Somerville and Wellman (1979) found a developmental dip in strategic memory use.

There was an interaction between age group and test type for response bias. Children were more likely to respond no/not-seen-before for the item test than young adults but there were smaller age differences in response bias for the associative test. This shows that children preferred to reject items and this may have artificially reduced their item memory performance (and minimised the difference between item and associative memory performance). A forced choice test may have eliminated this effect and may therefore have produced an age group by test type interaction indicating significant associative deficits in children relative to young adults.

There were also some effects of counterbalancing in the data. Even though the study material was the same for all participants, the way test material was constructed had effects on relatedness (the difference in performance between pairs that were related and unrelated at study) and affected item and associative memory differently. There was also slight evidence of a test order effect where there was an advantage to whichever test was completed first (item or associative). This may be due to the recency effect and/or to interference from the first test on the second test.

Despite some inconsistencies in the data, it is ultimately apparent that if children suffer from associative deficits, then it is likely that such deficits are not as extreme as those observed in older adults. This indicates that the development of associative memory ability in children does not mirror the decline of associative memory ability in older adults. Therefore, the current results provide some evidence against the view that associative memory is more complex as it requires more time to develop in children and that it is this complexity which makes associative memory more susceptible to age-related degradation. The following chapter aims to simulate age-related degradation in young adults for item and associative memory to elucidate if associative memory is more reliant on attentional resources (i.e., is more complex/requires more processing) than item memory.

Chapter 8: Associative Deficits and Divided Attention

Age-related cognitive deficits are typically viewed in consideration of global declines in cognitive functioning. Chapter 1 showed that cognitive speed (Salthouse, 1996), working memory functionality (e.g., Craik, 2000; Craik & Byrd, 1982) and ability to inhibit irrelevant material (e.g., Hasher & Zacks, 1988) all decline with increasing age. These theories can be respectively viewed as global declines in processing speed, working memory capacity and inhibitory processes (Park, 2000). Age-related associative memory deficits, however, are not easily reconcilable with the notion of a global decline in memory functionality with age and the current chapter aims to explore this anomaly.

In general, the evidence in the literature suggests that there is a disassociation between the effect of age on associative/context memory and item/content memory. Age-related associative/context memory deficits are found to be reliably greater than item/content memory deficits (Old & Naveh-Benjamin, 2008a; Spencer & Raz, 1995). A critic of this view could argue that item memory formation and retrieval is easier than associative memory formation and retrieval: It can be seen in many studies, including those reported in this thesis, that even young adults perform better at item memory tests than associative memory tests (e.g., Bastin & Van der Linden, 2006; Brockmole, Parra, Della Sala, & Logie, 2008; Naveh-Benjamin, Guez, & Marom, 2003). It may therefore be the case that item memory formation is simply less effortful than associative memory formation (Troyer, Winocur, Craik, & Moscovitch, 1999) and with the reduced self-initiated processing required, shows smaller age deficits (Spencer & Raz, 1995). Evidence from neuroimaging is also consistent with this view as greater prefrontal cortex activity has been observed for associative than item memory during both encoding and retrieval (Cabeza, 2006) and prefrontal tasks typically show greater age deficits (West, 1996). In contrast, there is evidence against the hypothesis that general difficulty differences between item and associative tests can explain age-related associative deficits: Age-related associative deficits are present even when overall memory performance is comparable between young and older adults (e.g., Naveh-Benjamin, Guez, Kilb et al., 2004; Naveh-Benjamin et al., 2009).

Importantly, associative deficits are generally not found in young adults when they complete memory tests under divided attention alongside a secondary task (e.g., Naveh-Benjamin, Guez et al., 2003; Naveh-Benjamin, Guez, & Shulman, 2004). That is, when task relevant cognitive resources are reduced in young adults via experimental manipulation, associative memory performance suffers to the same degree as item memory performance. If associative memory formation was dependent on higher levels of concentration and additional control, it would be expected to see this memory performance more highly disrupted by a secondary task than item memory performance. The current chapter aims to investigate this apparent contradiction in the literature where on the one hand we see associative memory tasks as more effortful and difficult than item memory tasks but on the other, they are equally susceptible to manipulation of available cognitive resources. The primary emphasis of this chapter is to investigate how the manipulation of cognitive resources in young adults via divided attention impacts upon item and associative memory formation and retrieval.

Divided Attention in Young Adults

When young adults complete memory tasks under divided attention (DA), they generally show a reduction in memory performance compared to when the tasks are completed under full attention (FA). This effect has received interest in relation to ageing research because it is possible to reduce memory performance in young adults to levels that are seen in older adults (Craik, 2006). For example, Anderson, Craik and Naveh-Benjamin (1998) measured the effects of ageing and DA. Young and older adults completed cued and free recall memory tests under both FA and DA. In the DA conditions, a secondary task (identifying, via a button press, which one of four boxes presented on a screen had a shape within it) was presented during encoding or during retrieval of the words. When instructed to prioritise the secondary task, DA at encoding reduced young adults' memory performance to the same level as older adults' memory performance under FA. Also, when instructed to prioritise the memory task during encoding, young adults' secondary task reaction times dropped to the same level as older adults performing the secondary task under FA. Young adults under DA performing similarly to older adults under FA is in line with a limited resources approach (e.g., Craik, 2000). This has led to a body of ageing research focusing on DA in young and older adults (see McDowd & Shaw, 2000, for review). Additionally, older adults are generally affected more than young adults by DA (McDowd & Shaw, 2000; Verhaeghen, Steitz, Sliwinski, & Cerella, 2003), which provides evidence that DA taxes processes that older adults rely on.

The effects of DA are somewhat unexpected when applied to associative deficits because DA does not affect associative memory in young adults more than item memory, despite the fact that associative memory tasks appear to be more demanding than item memory tasks. For example, Naveh-Benjamin, Guez and Shulman (2004) tested the role of cognitive resources in relation to associative deficits. Young and older adults were required to remember word pairs in a typical test of associative deficits – memory for the individual words represented item memory and memory for the pairings represented associative memory. The young group completed an additional condition where they had to perform a secondary task (digit monitoring) to divide attention during encoding. Typical age-related associative deficits were observed when older adults were compared to young adults under FA: Older adults' performance was worse than young adults for both item memory and associative memory but to a greater extent in the latter. Age-related associative deficits were also present when the older adults were compared to the young adults' performance under DA – there was no overall difference in performance due to age, but older adults had disproportionately lower performance on the associative test compared to young adults under DA. Dividing the attention of the young group affected both item and associative test performance equally; therefore reduced cognitive resources during the DA condition did not produce associative deficits in the young participants.

A search of the literature uncovered 23 experiments across 11 studies with young adults that had both FA and DA conditions as well as item and associative memory tests (Table 18 shows a summary of how DA affected item and associative memory performance in young adults). It can be seen that all but four of the 23 experiments showed a similar decrease in item and associative memory performance under DA compared to FA. There is no clear pattern between the studies showing associative deficits under DA, although three out of four of these experiments had divided attention during both encoding and retrieval (Troyer & Craik, 2000, Experiment 3; Troyer et al., 1999, Experiments 1 and 2). It must be noted, however, that for Troyer and Craik's (2000) Experiment 3, there was no pure FA measure because DA at encoding and retrieval was compared to DA at encoding only.

It can also be seen from Table 18 that the majority of the experiments only used DA during the encoding period. This is probably because DA at retrieval is not as reliable: Experiments have shown that dividing attention at encoding has a detrimental effect on memory performance but that dividing attention during retrieval only has a minimal effect on memory performance (e.g., Anderson et al., 1998; Baddeley, Vivien, Eldridge, & Thompson, 1984). However, there are also studies showing significant memory disruption when participants performed a secondary task during retrieval (e.g., Fernandes & Moscovitch, 2000; Fernandes & Moscovitch, 2002; Park, Smith, Dudley, & Lafronza, 1989). As there is a lack of DA studies that divide attention at retrieval, DA at retrieval will be examined further in the current study.

Table 18

Summary of Studies with Item and Associative Memory Measures Under Both Full and Divided Attention (FA and DA) Conditions

Study	Significant DA Costs			Memory Stimuli	Secondary Task	DA Period
	Item	Associative	Interaction[a]			
Castel and Craik (2003)						
Experiment 1a	Y	Y	Y	Unrelated word pairs	Digit monitoring	Encoding
Experiment 1b	N	N	N	Unrelated word pairs	Digit monitoring	Encoding and retrieval
Craik, Luo, and Sakuta (2010)						
Experiments 1	Y	Y	N	Words and pictures of scenes	Digit monitoring	Encoding
Experiments 2	Y	Y	N	Words and pictures of scenes	Digit monitoring	Encoding
Kersten and Earles (2010)						
Experiment 1	Y	Y	N	Actors and Actions	Judgement about a description	Encoding
Experiment 1	N	N	N	Actors and Actions	Judgement about a description	Retrieval
Kilb and Naveh-Benjamin (2007)						
Experiment 1	Y	Y	N	Unrelated word pairs	Tone identification	Encoding
Experiment 2	Y	Y	N	Unrelated word pairs	Tone identification	Encoding
Naveh-Benjamin, Guez, Kilb et al. (2004)	Y	Y	N	Names and faces	Tone identification	Encoding
Naveh-Benjamin, Guez et al. (2003)						
Experiment 1	Y	Y	N	Word-nonword pairs	Tone identification	Encoding
Experiment 2	Y	Y	N	Unrelated word pairs (auditory presentation)	Visual identification	Encoding
Experiment 3	Y	Y	N	Words and fonts	Visual identification	Encoding

Experiment 5	Y	Y	N	Related and unrelated word pairs	Visual identification	Encoding
Naveh-Benjamin, Guez, and Shulman (2004)	Y	Y	N	Unrelated word pairs	Digit monitoring	Encoding
Naveh-Benjamin, Hussain, Guez, & Bar-On (2003)						
Experiment 1	Y	Y	N	Unrelated pictures	Digit monitoring	Encoding
Experiment 2	Y	Y	N	Related word pairs	Tone identification	Encoding
Experiment 2	Y	Y	N	Unrelated word pairs	Tone identification	Encoding
Troyer and Craik (2000)						
Experiment 1	Y	Y	N	Words, colour of word background, order of words	Digit monitoring	Encoding
Experiment 2	Y	Y	N	Words, colour of word background, order of words	Digit monitoring	Retrieval
Experiment 3	N	Y[b]	Y[b]	Words, colour of word background, order of words	Digit monitoring	Encoding and retrieval[c]
Troyer et al. (1999)						
Experiment 1	Y	Y	Y	Auditory words and voice used	Finger tapping and visual identification	Encoding and retrieval
Experiment 2	Y	Y	Y	Auditory words and ear of headphones played through	Finger tapping and visual identification	Encoding and retrieval
Wang, Dew, and Giovanello (2010)	Y	Y	N	Unrelated word pairs	Prospective memory task – press button when a study word is an animal	Encoding

Note. All DA costs are for young adults.
[a]Interaction indicates a significantly greater DA cost to associative than item memory. [b]Associative drop and interaction only for extrinsic context (i.e., order of words) and not for intrinsic context (i.e., colour of background). [c]There was no pure FA condition as DA at encoding and retrieval was compared to DA at encoding only.

It may be the case that associative deficits in older adults arise primarily from deficits during the retrieval process. Cohn, Emrich and Moscovitch (2008) strongly argued in favour of this view – they found that associative deficits in older adults were greater when retrieval was more demanding. They hypothesised that associative memory was more reliant on recollection than item memory, which is based more on familiarity. Other studies have reached similar conclusions (e.g., M. Healy et al., 2005; Old & Naveh-Benjamin, 2008a). As was shown in Chapter 1, it is well established in the literature that recognition tests yield smaller age differences than recall tests (e.g., Craik & McDowd, 1987; Light et al., 2000; Naveh-Benjamin, 2000; Schonfield & Robertson, 1966). This led to a dual process theory of memory where recognition/familiarity is seen as a different process to recall, with the latter more susceptible to the ageing process. This therefore leads to a dual process explanation of associative deficits in older adults: Recall is impaired in older adults and therefore so is associative memory. This explanation of associative deficits is based entirely on the retrieval period.

A further point to consider is that age deficits in associative recognition tests are often driven by increased false alarms to lures whilst endorsement of seen-before associations remains relatively intact (e.g., Castel & Craik, 2003; Healy et al., 2005). This means that older adults have formed associative memories but that they experience difficulty using recollection to reject lures. Therefore, this provides more evidence that encoding is intact in older adults and that it is confusion at retrieval that causes the age-related associative deficits observed. In everyday life, older adults often experience difficulty retrieving information that they encoded many years ago (e.g., names of famous people or friends), even though such material was successfully encoded at the time (Craik, 2006). Finally, Naveh-Benjamin et al.

(2007) found that encouraging participants to use encoding strategies reduced age-related associative deficits but encouraging participants to use encoding *and* retrieval strategies almost eliminated age-related associative deficits in older adults.

In a study that only assessed associative memory, Naveh-Benjamin et al. (2005, Exp. 1) demonstrated that DA during encoding reduced associative memory in young and older adults equally. In Naveh-Benjamin et al.'s (2005) Experiment 2, young and older adults completed an associative memory task under FA and under DA during recall. Young adults' recall performance was unaffected by dividing attention but older adults showed reduced memory performance with the presence of the secondary task. This evidence suggests that older adults require more cognitive resources during recall. Naveh-Benjamin et al. (2005) also showed that performance on the secondary task dropped more for older adults than young adults during recall, especially when older adults were instructed to use memory strategies. The robust memory performance in young adults with DA during recall could be due to the nature of the secondary task, which may have tapped different resources to the memory task. It has been observed that there is little consensus in the literature about the best parameters to use for dividing attention (McDowd & Shaw, 2000). Several studies by Fernandes and colleagues have found that dividing attention during recall is more successful when the secondary task involves the same type of information as the material recalled (e.g., Fernandes & Moscovitch, 2002; Fernandes & Moscovitch, 2003; Fernandes, Moscovitch, Ziegler, & Grady, 2005). Fernandes and Moscovitch (2002) argued that the secondary task is more disruptive to memory if it is of a similar nature to the memory task. They refer to Baddeley and Hitch's (1974) dissociation between visual and phonetic information in working memory as a possible explanation for their result. Naveh-Benjamin et al.'s (2005) Experiment 2 used a visual tracking based secondary task during word recall so the lack of memory disruption in young adults could be due to separate demands from the concurrent tasks. In Table 18 it can be seen that all but two of the studies (Kersten & Earles, 2010; Wang et al., 2010) used basic perceptual secondary tasks. It may therefore be the case that the secondary tasks were not sufficiently disruptive to the more complex processes that are responsible for associative memory performance.

Experiment 8

The current study aimed to answer two questions: 1. Can a secondary task disrupt associative memory more than item memory when the secondary task is of a similar nature to the memory test? 2. Is there a differential effect of dividing attention at encoding, retrieval, or at both encoding and retrieval?

Firstly, the secondary task was chosen to maximise its chance of disrupting memory during retrieval and to also disrupt frontal activity. It used animacy judgements where participants had to decide if a given word corresponded to a living or a non-living thing. This task has been successfully used to disrupt memory performance by dividing attention during retrieval (Fernandes & Moscovitch, 2002). The task requires semantic, word-related judgements that are theoretically more disruptive to word memory performance than a basic perceptual or numerical task (as was used in the majority of literature testing associative deficits under DA). In a neuroimaging study, Kapur et al. (1994) found that when participants completed animacy judgements, there was increased activity in the left inferior prefrontal cortex compared to a condition where participants had to make a basic perceptual response (judging if a word contained the letter ‘a’). This demonstrates that the task is related to frontal activity and using it as a secondary task means that it should interfere with frontal processes. Dividing attention with this task is therefore more likely to simulate deficits seen in older adults who typically show prefrontal deficits (e.g., West, 1996). Also, Kapur et al. (1994) found that later memory performance was higher for words used in the animacy task than for words used in the basic perceptual task. This indicates that an animacy task probably uses processes related to memory performance and is therefore more likely to disrupt performance on the primary memory task.

Secondly, the current study compares the effect of DA versus FA on both item and associative memory as within participants factors using three different between participants DA conditions: DA at encoding, DA at retrieval, and DA at both encoding and retrieval together. This should establish if associative deficits are driven differentially by encoding and retrieval processes.

Method

Participants. Seventy-two participants took part in the experiment, 24 in each of the three conditions. Young adults were either paid £5 for participation or were given course credit. The 24 older adults from the words condition of Experiment 4 reported in Chapter 5 were also included for comparison. Table 19 shows the age, education and gender of the participants who took part in the study.

Table 19

Age, Education and Gender of Participants for the Three Experimental Conditions and for the Older Adults From Chapter 5

Condition	Age Range	Age	Education (years)	Number of Females
		M (SD)	*M (SD)*	
DA at encoding	18-28	20.3 (2.2)	14.5 (1.2)	14/24
DA at retrieval	18-25	19.2 (1.4)	14.0 (0.8)	22/24
DA at encoding and retrieval	18-30	20.8 (2.6)	15.1 (1.5)	17/24
FA older adults	65-85	75.3 (6.3)	13.0 (3.9)	15/24

Note. DA = Divided Attention, FA = Full Attention

Materials. For the measure of item and associative memory, six lists of unrelated word pairs were used. The lists were the exact same lists as used for the words condition in Experiment 4 (see the materials section of Experiment 4 for details of stimulus characteristics). Each list contained 34 word pairs in the memory test (including 2 pairs at the start and 2 pairs at the end that were used as buffers and were not tested). This was followed by item and associative memory tests. The item memory test consisted of 20 old words from the original memory set and 20 new, previously unseen, words. The associative memory test consisted of 10 intact and 10 rearranged pairs from the original memory set. For the recombined trials of the associative tests, words originally displayed on the left remained on the left and words originally displayed on the right remained on the right. Words appearing in the item test did not appear in the associative test. Words were displayed on a computer monitor in font size 40 point with a height corresponding to roughly 1°

viewing angle at a distance of 60 cm. There was also a practice version of the task with three word pairs containing words different to those used in any of the lists.

For the secondary task, three lists of 50 words were used for the animacy judgement task. In each list, half of the words corresponded to living things (e.g., dolphin, puppy, rattlesnake) and half to non-living objects (e.g., computer, glasses, luggage). The words were acquired from Fernandes' laboratory where experiments had been previously conducted that disrupted memory performance when using similar lists as secondary tasks during retrieval (e.g., Fernandes & Moscovitch, 2003; Skinner & Fernandes, 2008). Ten of the words for the secondary task were replaced because they were also used in the main recognition memory task. The English lexicon project database (Balota et al., 2007) was used to assess certain characteristics of the words that were used. The words varied from 3-11 letters in length (*M* = 5.54, *SD* = 1.85) and they occurred with an average frequency of 8.26 (*SD* = 1.64, range = 4.19 – 12.44), using log HAL frequency (Lund & Burgess, 1996). The words were recorded digitally and each was normalised for equivalent maximum amplitude. During the experiment, the words were played through powered speakers attached to a laptop computer. Eight extra words were also used to make a practice version of the task.

Procedure. Participants were informed that they were going to be required to complete a recognition memory test under two conditions, one normal (FA) and another alongside a secondary task (DA). At the beginning of the session, participants were given a short practice at the secondary task on its own (six trials). They were asked to listen to the secondary task words played through the speakers and after each word respond verbally 'living' or 'non-living'. The words were played at a rate of one word every 4 s and this rate was fixed regardless of whether a response was made or not. Immediately after each response, the experimenter pressed a button to record their choice and a computer recorded the time of each response in relation to the initial onset of each word. This provided a measure of response times to each word and is a technique used in much prior research (see Maylor, Watson, & Hartley, 2011). In addition, reliability was maintained as the same experimenter tested every participant. After the practice block, the secondary task was completed with one of the lists of 50 words in the same manner. This provided a baseline measure of secondary task performance in isolation without the memory test at the same time.

Participants then practised the recognition memory test. They were shown three word pairs at a rate of one pair every 4 s. Then, before they were tested on their word memory, there was a 1-minute delay. This was filled with 30 s of backwards counting in threes from 300 followed by 30 s of responding to the secondary task. Participants completed a two-trial item recognition test where they had to respond via button presses (keys ‘j ‘and ‘f’) on a keyboard as to whether each item was old (presented in the original memory set) or new. They also completed a two-trial associative memory test where they were presented with word pairs that were either intact or recombined; again they were asked to respond via a button press as to whether they thought they had seen the stimuli paired together before (‘j’) or not (‘f’). A written description of the responses was left on the desk so that participants did not have to remember the buttons. The order of the tests (item then associative or associative then item) was counterbalanced across participants in each condition. After the first practice, the participants were given another practice test, this time with the secondary task running at the same time so that they would practice completing the memory test at the same time as responding to the secondary task.

For the DA at encoding condition, the secondary task was performed during the presentation of the memory set. For the DA at retrieval condition, the secondary task was performed during the item and associative memory tests. For the DA at encoding and retrieval condition, the secondary task was performed during both of the aforementioned periods. Participants were instructed to divide their attention equally between the two tasks.

After the practice sessions, the main memory tasks were completed in the same way as described for the practice sessions, once under FA and once under DA. The DA test varied according to the condition to which the participant was assigned. There was a short rest between the FA and DA memory tests.

Counterbalancing and randomisation. Participants were randomly assigned to one of the three secondary task lists. This list would be the one they performed without the concurrent memory test. The same list was then used in the practice tests and during the 1-minute delay between encoding and retrieval in the main memory tests. The remaining two secondary task lists were then combined to be used in the DA memory test. Individual secondary task words were presented in random order. For the primary memory task, there were six sets of words corresponding to six study-test lists. Within each list, words and pairs were manually assigned to the memory set, the item test or the associative test. The presentation of individual words and pairs was randomised within the memory set, the item test and the associative test.

For counterbalancing, the six different study-test lists were grouped into three pairs that were yoked to the three secondary task lists. This produced three levels based on *stimulus type*. The order of use of each of the two study-test lists was counterbalanced to give two levels of *list order*. The order of memory test was also counterbalanced so that a given participant would always have the item test before the associative test or vice versa, giving two levels of *test order*. Finally, the order of the DA test was counterbalanced so that half of the participants had DA then FA, and half FA then DA, giving two levels of *DA position*. This produced a 3 x 2 x 2 x 2 design with 24 counterbalancing combinations; one was given to each participant within the three different DA conditions (24 participants in each group).

Results

Memory performance. Performance on the recognition memory tasks was ascertained by calculating hit rates minus false alarm rates, d' and ln(β) as described in Chapter 5. The analysis presented is based on hit rates minus false alarm rates; d' data were also analysed and any notable differences in results between the measures are reported.

Initially, tests were carried out to ascertain effects of the different counterbalancing conditions. The effects of the four counterbalancing factors (stimulus type, list order, test order, and DA position) were assessed one at a time and were entered as a between subjects variable in a 2 (Attention: FA, DA) x 2 (Memory test: item, associative) repeated measures ANOVA based on the hits minus false alarms measure. The ANOVAs were conducted separately for each of the three DA conditions. There was no main effect of stimulus type, list order or test order, or any interactions with stimulus type, list order or test order, for any DA condition. However, there was evidence of an influence of DA position. Although there was no main effect of DA position for any DA condition, there was a triple interaction between attention, test and DA position for the DA at retrieval condition, $F(1, 22) = 5.12$, $MSE = 0.01$, $p < .05$. No other interactions were present involving DA position for any DA conditions. The effect of DA position was therefore considered in the following comparisons, but the other counterbalancing criteria were not used as factors in any further analysis.[33]

To assess the overall pattern of the data for the young adults, hit rates minus false alarm rates were entered into a 2 (Attention: FA, DA) x 2 (Memory test: item, associative) x 3 (DA condition: DA at encoding, DA at retrieval, DA at encoding and retrieval) mixed ANOVA (see Figure 26 for means). There was a main effect of attention, with performance under FA exceeding performance under DA, $F(1, 69) = 53.73$, $MSE = 0.03$, $p < .001$. There was also a main effect of memory test with item test performance exceeding associative test performance, $F(1, 69) = 17.34$, $MSE = 0.04$, $p < .001$. There was no main effect of DA condition, $F < 1$. Crucially, with regards to associative deficits under DA, there was no interaction between attention and memory test, $F(1, 69) = 1.90$, $MSE = 0.03$, *ns*. This means that DA did not affect associative memory significantly more than item memory. No other interactions were present (all remaining *F*s < 1.18). The results were qualitatively identical with d'.

[33] With d' the effects of counterbalancing were qualitatively identical and the triple interaction between attention, test and DA position for the DA at retrieval condition was significant, $F(1, 22) = 4.79$, $MSE = 0.17$, $p < .05$.

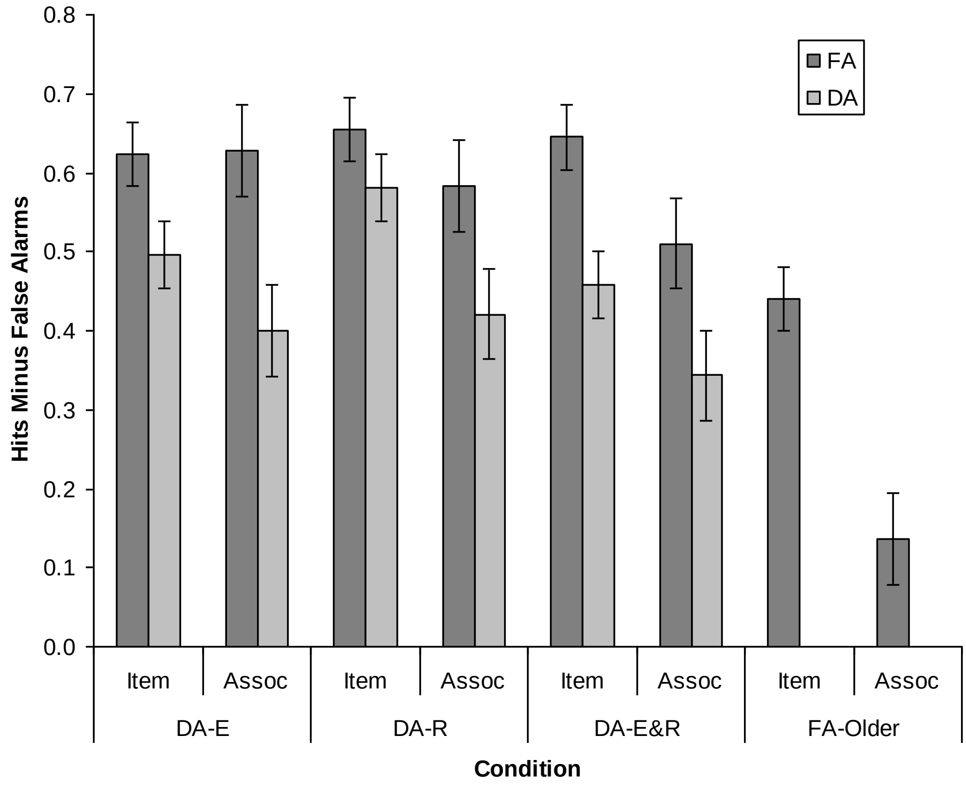

Figure 26. Hits minus false alarm performance for item and associative memory tests across four conditions: DA-R, divided attention (DA) during retrieval; DA-E, DA during encoding; DA-E&R, DA during encoding and retrieval; FA-Older, full attention older adults words performance from Experiment 4 . Error bars are ± 1*SE*.

The analysis was repeated adding DA position counterbalancing as a between subjects factor. There were no interactions involving DA position although there was a marginal effect of DA position, $F(1, 66) = 3.77$, $MSE = 0.19$, $p = .056$, with performance overall being higher for participants who received the FA memory test before the DA memory test than for participants who received the reverse (.58 vs. .48, respectively). Again the results were qualitatively identical with d' and the effect of DA position was also marginal, $F(1, 66) = 3.59$, $MSE = 2.48$, $p = .063$.

Age-related associative deficits. To test if age-related associative deficits were present with the current data, the FA data for all three DA conditions together were compared to the older adults under FA for words from Experiment 4. A 2 (Age: young, older) x 2 (Memory test: item, associative) mixed ANOVA was conducted on the hit rates minus false alarm rates. There was a main effect of age, $F(1, 94) = 38.77$, $MSE = 0.10$, $p < .001$, with young participants ($M = .61$, $SD = .26$) outperforming older participants ($M = .29$, $SD = .16$). There was also a main effect of test, $F(1, 94) = 50.26$, $MSE = 0.03$, $p < .001$, with item test performance ($M = .54$, $SD = .22$) exceeding associative test performance ($M = .36$, $SD = 0.34$). There was an interaction between age and memory test, $F(1, 94) = 20.52$, $MSE = 0.03$, $p < .001$, with a greater drop in performance from item to associative memory in older adults (.31) relative to young adults (.07), thus demonstrating a clear age-related associative deficit. All of the results were qualitatively identical with d'.

More importantly with regards to the current study, the same analysis was conducted using the performance of young adults under DA. There was a main effect of age, $F(1, 94) = 10.74$, $MSE = 0.09$, $p = .001$, with young participants ($M = .45$, $SD = .27$) outperforming older participants ($M = .29$, $SD = .16$). There was also a main effect of test, $F(1, 94) = 48.97$, $MSE = 0.03$, $p < .001$, with item test performance ($M = .48$, $SD = .21$) exceeding associative test performance ($M =.26$, $SD = .30$). Crucially, there was an interaction between age and memory test, $F(1, 94) = 8.77$, $MSE = 0.03$, $p < .01$, with a greater drop in performance from item to associative memory in older adults (.31) relative to young adults (.12). Follow up analysis revealed that for the item test there was no age difference,[34], $t(64.71) = 1.81$, $p = .08$, but for the associative test there was, $t(71.72) = 5.06$, $p < .001$. It can also be seen in Figure 26 that item memory performance for young adults under DA in the DA at encoding and retrieval condition was nearly identical to that of FA older adults ($t < 1$); however, associative memory was still lower in older adults compared to the young adults in this DA condition.[34] $t(36.75) = 2.99$, $p < .001$. This means that even

[34] Levene's test for equality of variances was significant so equal variances were not assumed.

though DA successfully dropped young adults' item memory to that of older adults, age-related associative deficits were still present. All of the results were qualitatively identical with d'.

Divided attention conditions. The three DA conditions were analysed separately to establish if there was a differential effect of DA position between them (see Appendix 6 for means). This would clarify if DA position affected the interaction between attention and memory test differently when attention was divided at different periods of the memory test. For DA at encoding, a 2 (Attention: FA, DA) x 2 (Memory test: item, associative) x 2 (DA position: DA then FA, FA then DA) mixed ANOVA was conducted on the hit rates minus false alarm rates. There was a main effect of attention, $F(1, 22) = 27.30$, $MSE = 0.03$, $p < .001$, with performance under FA exceeding performance under DA. There was no main effect of memory test, $F(1, 22) = 1.40$, $MSE = 0.04$, *ns*. There was no main effect of DA position, $F(1, 22) = 2.85$, $MSE = 0.18$, *ns*. There was a marginal interaction between memory test and DA position, $F(1, 22) = 2.88$, $MSE = 0.04$, $p = .10$, because associative memory performance was very slightly higher than item memory performance for the DA-FA group but item memory performance was higher than associative memory performance for the FA-DA group. This is perhaps the source of the lack of a main effect of memory test. There were no other interactions.

The same analysis was conducted for the DA at retrieval condition. There was a main effect of attention, $F(1, 22) = 11.33$, $MSE = 0.03$, $p < .01$, with performance under FA exceeding performance under DA. There was also a main effect of memory test, $F(1, 22) = 7.93$, $MSE = 0.04$, $p = .01$, with item test performance exceeding associative test performance. There was no main effect of DA position, $F(1, 22) = 1.16$, $MSE = 0.17$, *ns*. Interestingly, there was a marginal interaction between attention and memory test, $F(1, 22) = 3.37$, $MSE = 0.01$, $p = .08$, because DA reduced associative memory performance more than item memory performance. There was a three-way interaction between attention, memory test and DA position, $F(1, 22) = 5.12$, $MSE = 0.01$, $p < .05$, because DA did produce a greater drop in associative memory performance than item memory performance for participants who had the DA test before the FA test (DA-FA) but not for participants who had the DA test after the FA (FA-DA) test. There were no other interactions in the data.

Finally, the same analysis was conducted for the DA at encoding and retrieval condition. There was a main effect of attention, $F(1, 22) = 16.64$, $MSE = 0.05$, $p < .001$, with performance under FA exceeding performance under DA. There was also a main effect of memory test, $F(1, 22) = 10.11$, $MSE = 0.04$, $p < .01$, with item test performance exceeding associative test performance. There was no main effect of DA position, $F < 1$. None of the interactions was significant.

To summarise, there was some evidence that DA was more detrimental to associative than item memory when attention was divided at retrieval only: Thus if the participants completed the memory test under DA before the FA memory test, DA produced associative deficits. This DA position effect was not apparent in the other two DA conditions although for DA at encoding a numerically similar pattern was found. To remove any effects of DA position, half of the data were removed so that the following analysis was conducted on only the first memory test that participants received. This meant that attention was now a between subjects factor. A 2 (Attention: FA, DA) x 3 (DA condition: DA at encoding, DA at retrieval, DA at encoding and retrieval) x 2 (Memory test: item, associative) ANOVA was conducted on the hit rates minus false alarm rates for the first memory test taken. There was a main effect of attention, $F(1, 66) = 22.55$, $MSE = 0.10$, $p < .001$, with performance under FA exceeding performance under DA. There was also a main effect of memory test, $F(1, 66) = 12.47$, $MSE = 0.04$, $p < .001$, with item test performance exceeding associative test performance. There was no main effect of DA condition, $F < 1$. Crucially, there was no interaction between attention and memory test, $F(1, 66) = 2.42$, $MSE = 0.04$, $p = .13$. None of the other interactions was significant.

Reaction times and secondary task performance. Secondary task responses were analysed to assess how much the primary task reduced accuracy and speed at the secondary task. This would provide a measure of how much resources were used by the primary task. Secondary task performance when it was performed alone was used as a baseline comparison. The last responses were excluded from the secondary task performed during encoding and from the secondary tasks performed during the item and associative recognition tests. This was because the primary memory phase/tests were not always running when the last responses were made (i.e., the last responses were not always made under DA).

Accuracy on the secondary task was very high at over 84% in all conditions. The DA conditions were analysed separately because not all of them had the same secondary task measures. For the DA at encoding condition, accuracy at the secondary task was no different during encoding ($M = .95$, $SD = .08$) compared to baseline ($M = .94$, $SD = .07$), $t < 1$. For the DA at retrieval condition, accuracy at the secondary task during the item memory test ($M = .89$, $SD =.13$) was marginally worse than baseline ($M = .94$, $SD =.09$), $t(23) = 1.94$, $p = .07$. Accuracy at the secondary task during the associative memory test ($M = .84$, $SD =.15$) was significantly worse than baseline, $t(23) = 3.86$, $p < .001$. For the DA at encoding and retrieval condition, accuracy at the secondary task during encoding ($M = .95$, *SD*

=.05) was also no different to baseline (*M* = .95, *SD* =.06), *t* < 1. Accuracy at the secondary task during the item test (*M* = .93, *SD* =.08) was no worse than baseline, *t*(23) = 1.37, *ns*. Accuracy at the secondary task during the associative test (*M* = .91, *SD* =.10) was significantly worse than baseline, *t*(23) = 2.27, *p* < .05.

Another measure of secondary task costs and now also primary task costs was obtained from the reaction time data. Reaction times were computed on a proportional basis such that DA reaction times were represented as a proportion of FA reaction times (e.g., Lindenberger, Marsiske, & Baltes, 2000). This would account for general differences in speed between participants. The following equation was used to measure dual task costs for each participant:

$$\text{Dual Task Costs} = 1 - \frac{FA_{RT} - DA_{RT}}{FA_{RT}} \quad (7)$$

Where FA_{RT} = FA reaction time and DA_{RT} = DA reaction time (see Figure 27 for dual task costs for the primary and secondary task reaction times). Again the three DA conditions were analysed separately. One sample *t*-tests measured if dual task costs were significantly above 1 (i.e., the presence of dual task costs).

For the DA at encoding condition, reaction times for the secondary task were significantly slower when it was performed during encoding, *t*(23) = 5.49, *p* < .001. Interestingly, reaction times for the primary item recognition test were also slower when the word pairs had been encoded under DA compared to FA, *t*(23) = 2.83, *p* < .01, even though the test phase was completed under FA for both DA and FA encoding. Primary task associative recognition performance was not significantly changed by DA, *t* < 1.

For the DA at retrieval condition, there was a significant increase in reaction time for the secondary task when it was performed at the same time as the item recognition test, *t*(23) = 7.83, *p* < .001; likewise, compared to the FA memory test, reactions to the items in the primary task item memory test were also significantly slower, $t(23) = 9.13$, $p < .001$. The same pattern was observed for the associative memory test; secondary task responses were slower, $t(23) = 10.45$, $p < .001$, and the primary task associative recognition responses were slower, $t(23) = 5.85$, $p < .001$.

For the DA at encoding and retrieval condition, there was a significant increase in reaction times for the secondary task when it was performed at the same time as encoding, $t(23) = 5.44$, $p < .001$, the item memory test, $t(23) = 9.55$, $p < .001$, and the associative memory test, $t(23) = 11.17$, $p < .001$. Similarly, the reaction times to the primary item memory test were slower under DA, $t(23) = 9.50$, $p < .001$, and so were reaction times to the primary associative memory test, $t(23) = 5.84$, $p < .001$.

The amount of DA slowing on the secondary task during the item test was compared to the amount of DA slowing during the associative memory tests. For the DA at retrieval condition, reaction times on the secondary task were slowed significantly less when it was performed during the item test compared to when it was performed during the associative test, $t(23) = 2.26$, $p < .05$. For DA at both encoding and retrieval, reaction times at the secondary task were slowed similarly under DA for the item and associative tests, $t < 1$. The slowing of the secondary task due to DA was greater when it was performed during the item test compared to when it was performed during encoding, and similarly for the associative test compared to encoding ($t(23) = 8.04$, $p < .001$, and $t(23) = 7.96$, $p < .001$, respectively).This shows that DA at retrieval potentially caused more dual task conflict than DA at encoding, presumably because participants were required to respond to two tests at once.

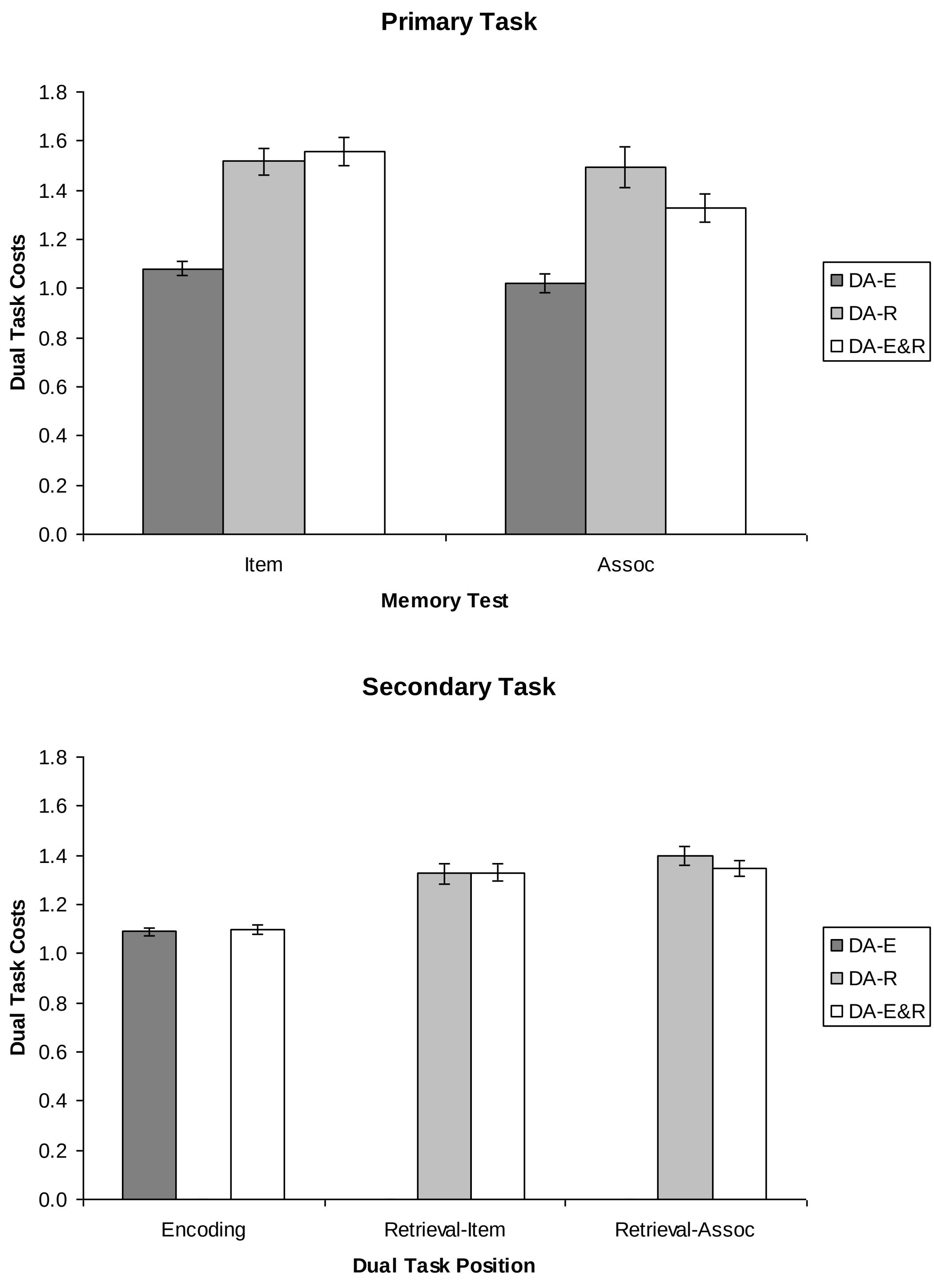

Figure 27. Dual task costs to reaction times for the three conditions DA at encoding (DA-E), DA at retrieval (DA-R) and DA at both encoding and retrieval (DA-E&R). DA costs for primary memory task (top) reported for item (Item) and associative (Assoc) memory tests. DA costs reported for secondary task (bottom) performed during encoding (Encoding), during item test retrieval (Retrieval-Item) and during associative test retrieval (Retrieval-Assoc). Error bars are ± 1*SE*.

A 3 (DA condition: DA at encoding, DA at retrieval, DA at encoding and retrieval) x 2 (Dual task cost: primary item memory test slowing, primary associative memory test slowing) mixed ANOVA was conducted on the primary memory tests dual task costs data. There was a main effect of dual task cost, $F(1, 69) = 9.03$, $MSE = 0.04$, $p < .01$, with costs higher for item (1.39) than for associative (1.28) memory tests. There was a main effect of DA condition, $F(2, 69) = 26.46$, $MSE = 0.11$ $p < .001$, with dual task costs higher for DA during retrieval (1.51) and DA during encoding and retrieval (1.44) than for DA during encoding (1.05). This pattern is most likely driven by the fact that, for the DA at encoding condition, participants did not have to respond to the secondary task during retrieval, therefore DA did not disrupt the primary memory task performance as much. There was also an interaction between DA condition and dual task cost, $F(2, 69) = 3.40$, $MSE = 0.04$, $p < .05$, because although item memory dual task costs were higher than associative memory dual task costs in all conditions, the difference was much greater for the DA at encoding and retrieval condition. This was confirmed by *t*-tests. The amount of slowing of reactions to the primary memory test under DA was not significantly different between the item and associative memory tests for the DA at encoding or for DA at retrieval conditions ($t(23) = 1.46$, $t < 1$, *ns*, respectively). However for the DA at encoding and retrieval condition, reactions to the primary task item memory test were slowed significantly more under DA than reactions to the associative memory test, $t(23) = 4.66$, $p < .001$.

Response bias. The data were also analysed to assess response bias for the primary memory task. Table 20 shows the mean response bias for each of the DA conditions: ln(β) provides a negative value for bias towards ‘yes/seen-before’ responses and a positive value for bias towards ‘no/not-seen-before’ responses. In general there is a response bias towards ‘no/not-seen-before’. A 2 (Attention: FA, DA) x 2 (Memory test: item, associative) x 3 (DA condition: DA at encoding, DA at retrieval, DA at encoding and retrieval) mixed ANOVA was conducted on the response bias measure ln(β) for the primary item and associative memory test responses. There was no main effect of attention, $F < 1$. There was a main effect of test, $F(1, 69) = 5.83$, $MSE = 0.43$, $p < .05$, because although both item and associative tests had a bias towards ‘no/not-seen-before’ responses, the bias was greater for the item test. There was no main effect of DA condition, $F(2, 69) = 1.94$, $MSE = 0.54$, *ns*, but there was an interaction between memory test and DA condition, $F(2, 69) = 3.15$, $MSE = 0.43$, $p < .05$. This is because the associative test showed a minimal response bias across all DA conditions but the item test showed a strong bias towards ‘no/not-seen-before’ for the DA at retrieval condition but a minimal response bias for the DA at encoding and the DA at encoding and retrieval conditions. The DA at retrieval condition showed a bias that was similar to FA older adults. However, this was the same for both FA and DA tests so was unlikely to be related to the manipulation of attention.

Table 20

Response Bias ln(β) for the Three DA Conditions and for FA Older Adults

		Test			
DA Condition	Attention	Item		Associative	
		M	*SD*	*M*	*SD*
DA-E	FA	0.25	0.90	0.09	0.53
DA-R		0.62***	0.55	0.07	0.67
DA-E&R		0.15	0.78	0.07	0.54
DA-E	DA	0.16	0.67	0.22*	0.46
DA-R		0.54**	0.72	0.16[*]	0.44
DA-E&R		0.21	0.62	0.19*	0.38
FA-Older		0.56***	0.61	0.04	0.21

Note. DA-E, DA at Encoding Only; DA-R, DA at Retrieval Only; DA-E&R, DA at Both Encoding and Retrieval. Significance values correspond to significant differences from zero.
[*] $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Discussion

In general, the manipulation of attention was successful as performance dropped during divided attention in all three conditions: DA at encoding, DA at retrieval and DA at both encoding and retrieval. The experiment also successfully disrupted memory at retrieval, which has been difficult to achieve in the past with perceptual and numerically based secondary tasks (e.g., Anderson et al., 1998; Baddeley et al., 1984). The extent of memory disruption at retrieval by the animacy judgement secondary task was even equivalent to the disruption caused during encoding. Therefore, the current results add further support to the use of this word-specific secondary task as a method of reducing word memory performance at all stages of a memory test (Fernandes & Moscovitch, 2000). Contrary to our expectations, the use of a word-specific secondary task did not differentially disrupt item and associative memory. Nor was there a differential effect of DA for DA during encoding, retrieval or both together. This is in line with the majority of studies that divided attention in tests of item and associative memory, where item and associative memory showed equivalent declines under DA (e.g., Craik et al., 2010; Naveh-Benjamin, Guez, & Shulman, 2004; see Table 18 for summary). When compared to older adults, the data from the young adults under DA still produced an age by memory test interaction, whereby older adults compared to young adults showed a disproportionately impaired associative memory relative to item memory. Overall the results are consistent with an age-specific associative deficit: Older adults show particular difficulty at forming associations, even when compared to a group of young adults whose performance is reduced by a concurrent, resource-demanding task. Again, in line with previous views, this suggests that age-related associative deficits occur independently of age-related declines in cognitive resources (Craik et al., 2010).

With regards to reaction times, DA significantly disrupted speed of response for both the primary and secondary tasks in all of the DA conditions.[35] When the secondary task was completed alongside memory retrieval, the primary memory recognition responses slowed. This is another measure of interference that was also largely consistent with the memory accuracy data: A slowing of response times corresponded to a reduction in memory performance. Some aspects of the reaction time data were inconsistent; for the DA at retrieval condition, secondary task response speed was disrupted significantly more during the associative test compared

[35] For the DA at encoding condition, the primary memory task associative test was slowed by DA at encoding compared to FA at encoding but not significantly.

to the item test. However, for the DA at encoding and retrieval condition, there was no difference in secondary task speed costs between item and associative memory tests. There was also an inconsistency in the primary memory test reaction data; both the DA at encoding and the DA at retrieval conditions data yielded no differences in reaction times between the primary item and associative recognition responses, but the DA at both encoding and retrieval condition showed smaller dual task costs for the associative recognition test than the item recognition test. Also, when attention was divided at encoding only, recognition responses to the primary item memory test were also slowed even though the test itself was completed under FA. Overall, the data did not provide evidence for a difference in dual task costs between item and associative memory.

Response bias was consistently conservative across all conditions, with participants biased to say that they had not encountered stimuli before. There was no main effect of response bias between FA and DA so the presence of the secondary task did not influence which response participants preferred in general. With regards to FA, for the DA at retrieval condition there was a significant response bias in the item test towards ‘no/not seen before’. This is inconsistent with the corresponding tests in the other two DA conditions, which showed no significant response bias (even though for FA, the conditions were identical). Under DA, all three conditions showed a response bias towards ‘no/not seen before’ on the associative test, but only the DA at retrieval condition showed this bias significantly for the item test. It is difficult to establish any clear differential effects between item and associative memory tests although there is some evidence that dividing attention provided a bias towards no/not-seen-before in the associative test more so than the item test.

The effect of the position of the DA memory test produced an interesting interaction. For DA at retrieval, if the DA task was completed first then DA reduced associative memory more than item memory. When the DA task was completed second, item and associative memory were similarly disrupted by dividing attention. This resulted in no net significant interaction between memory test and attention. A numerically similar but non-significant pattern was also found with the DA at encoding condition but not the DA at encoding and retrieval condition. Looking at the first test only for all three conditions (therefore removing any effects of the position of the DA test) yielded no greater drop due to DA in associative than item memory. It is possible that there was some effect of practice that caused this pattern of data although none of the studies listed in Table 18 reported any effects related to practice or counterbalancing of the DA task. Studies by Kramer and colleagues (Kramer, Larish, & Strayer, 1995; Kramer, Larish, Weber, & Bardell, 1999) showed that training can improve dual task performance so it may be that the associative test was more susceptible to disruption when participants were new to the procedure. Kramer, Larish, Weber and Bardell (1999) also showed that dual task costs, which were greater in older adults, were reduced relative to young adults with practice. It must also be noted that participants practised both the FA and DA conditions before completing the full tests so there is no reason to suspect that unfamiliarity with the test could have caused this effect.

One aspect of the young adults' results that is particularly interesting to consider is the additive effect of DA for the DA at encoding and the DA at retrieval conditions. It can be seen in Figure 26 that for item memory, the drop in memory performance from FA to DA in the DA at encoding and retrieval condition is very similar to the drop seen for the DA at encoding condition plus the drop seen for the

DA at retrieval condition. This is logical because the design of the DA at encoding and retrieval condition combines the other two conditions. What is interesting is that the same is not true for associative memory: The DA at encoding and retrieval condition shows a similar drop to the DA at encoding condition and the DA at retrieval condition individually (i.e., no additive effect of DA at both encoding and retrieval). This pattern in the data points towards an associative system in young adults that is somehow spared from the additive effects of DA at both encoding and retrieval. This therefore indicates a new type of inconsistency between item and associative memory performance. It is possible that associative memory is spared from additive effects of DA because with the DA at encoding and retrieval condition, both encoding and retrieval occur in the same dual-task environment. This is in line with the encoding specificity hypothesis where memory is improved when encoding and retrieval occur in the same context (Tulving & Thompson, 1973). It may also be the case that DA does not tax areas of the brain related to associative memory.

The effects of DA in the current study did not mimic the effects of age, contrary to hypotheses that view attentional resources as responsible for age deficits in memory (e.g., Craik, 1982). This is unusual because ageing and DA have been shown to have similar effects on pre-frontal activity, both resulting in lower activation of the left ventral prefrontal cortex (as reported in Craik et al., 2010). Furthermore, it has also been shown that pre-frontal activity is related to deeper processing and improved memory performance (Kapur et al., 1994). It may be the case that general age deficits in memory can be explained by frontal lobe status but that binding and associative memory are mediated by the medial temporal lobe (Cohn et al., 2008). It has been found that there is an age-related decrease in medial temporal lobe activity and that this area also has an impact on memory performance

(Dennis & Cabeza, 2008). Additionally, the medial temporal lobe has been linked to declarative/episodic memory, which requires associative memory formation (Fernandez & Fell, 2006). This indicates that the current study may have only disrupted frontal performance with DA, resulting in equal DA costs to item and associative memory, although contrary to these views Braver and West (2008) argue that age differences in frontal lobe status can explain age deficits in associative memory but not item memory. It remains for future research to establish a relation between age-related associative deficits and medial temporal lobe functionality.

Benjamin (2010) argued strongly against a specific age deficit in associative memory, hypothesising that global declines in memory fidelity may appear enhanced for associative memory measures. A key point he raises is that there are no known populations for whom associative memory is spared but item memory shows deficits. Benjamin postulates that this provides evidence against a qualitative distinction between item and associative memory. In his review, however, Benjamin does not explore the effects of divided attention and cites none of the studies in Table 18. The current study and the majority of previous research into associative memory deficits under DA suggest that a global decline in resources cannot account for age-related associative deficits. However, it is apparent that item and associative memory are both reliant on an underlying memory mechanism. In a meta-analysis, Old and Naveh-Benjamin (2008a) found a significant correlation between item memory and associative memory age differences ($r(88) = .39$, $p < .001$). They reported that across a range of experiments, when there was a large age deficit in item memory there was a correspondingly large age deficit in associative memory. Despite this relationship, there appears to be something extra required in the binding process and age-related associative memory deficits are consistently greater than item memory deficits.

Experiment 8 builds on existing research into the role of attention in item and associative memory. The results extend previous findings of equivalent disruption to item and associative memory from DA at encoding to DA at retrieval. Overall, they indicate that the age-related associative deficit cannot be accounted for by a global reduction in attentional resources with ageing. This implies that factors other than complexity and difficulty of associative memory tasks are responsible for the age-related associative deficit.

Chapter 9: Conclusion

This chapter summarises and integrates the results and conclusions from the previous chapters. The overall approach of the thesis was to assess resource-based hypotheses of cognitive ageing in terms of associative memory. The thesis explored the influence of support on age-related associative deficits where associative memory processing was encouraged by distinctiveness and the use of preexisting knowledge. A direct assessment of cognitive resources upon associative memory processes was also conducted with young adults in a divided attention study. The current chapter also considers differences between item and associative memory in relation to cognitive ageing in a more speculative manner. This leads on to discussion of future directions for research into the age-related associative deficit.

Overview of Findings

Distinctiveness. Chapter 3 measured the isolation effect in older adults. Tests were conducted to establish if a distinctive stimulus would be more easily remembered among homogenous control stimuli by older adults as is found in young adults. In the literature there were mixed results as to whether older adults did or did not show an isolation effect (see Chapter 3). Chapter 3 tested the hypothesis that age differences in the isolation effect are caused by age-related associative deficits. This is because the distinctiveness of an isolate depends on associating the isolated stimulus to the isolating factor (e.g., associating a word to its colour of presentation). Age differences in the isolation effect were explored in terms of modality and awareness of isolates, because modality and attention have been shown to influence the associative deficit in older adults.

Experiment 1 found that there were no age deficits in ability to show an isolation effect. Additionally, older adults produced an isolation effect that was no different in magnitude to the isolation effect in young adults. The results yielded different magnitudes of isolation effects from different types of isolation. Colour and position isolation did not yield strong isolation effects in young or older adults yet modality isolation was successful for both age groups. The age equivalence in the modality isolation effect was congruent with research showing minimal age-related associative deficits for modality information (Old & Naveh-Benjamin, 2008a). That is, older adults may have been able to show modality isolation because they were able to associate modality isolates to their isolating factor.

The second part of Chapter 3 assessed if age differences in the isolation effect could be due to age differences in awareness. The data from Bireta et al. (2008) were obtained. In their study older adults showed a smaller isolation effect than young adults on average. Their data were reanalysed to see if this could be explained by older adults taking longer to notice that some stimuli were isolated. Data from the first and second halves of Bireta et al.'s (2008) experiments were analysed to establish if older adults showed larger isolation effects in later trials once they were used to the experimental procedure (i.e., once they were familiar with isolates). There was no evidence that older adults took longer to notice isolates in the data as their isolation effect was similar in earlier and later trials.

Finally, in Experiment 2 of Chapter 3, the isolation effect was measured in young and older adults where the isolating factor itself was determined by associations between components of stimuli. For control stimuli two different types of information were presented together (i.e., two digits and two letters). Isolates were created by presenting uniform stimuli consisting of all digits or all letters. This meant that to notice isolates, participants would need to be aware of the structure of stimuli, which required associative memory. The results showed that even when the criteria for isolation depended heavily on associative memory, isolation effects were similar in young and older adults. Interestingly, because older adults produced an isolation effect, the isolation actually enhanced associative memory: Successful memory was determined by recognising which components of individual stimuli occurred together. This required memory for associations between items. Therefore, when isolation boosted memory it boosted associative memory. This indicates that the salience of isolates can provide environmental support, which aids associative memory in older adults.

Overall, the results from Chapter 3 did not indicate that age-related associative deficits influence the magnitude of the isolation effect: Given that age-related associative deficits are a reliable finding, there appeared to be no consistent link between these deficits and age differences in the isolation effect. Therefore, later chapters explored the associative deficit in different contexts.

Chapter 4 went on to explore the role of distinctiveness in associative memory in an applied context. Young and older adults were tested on their memory for distinctive faces in an experiment designed to simulate eye witness identification. Participants were shown a sequence of faces (some of which were distinct target faces) and then had to later identify those faces in lineups. Distinctive faces stand out in a lineup producing a biased test of memory and this is a problem for police forces conducting criminal investigations. To avoid this biased test, distinctive features were either replicated across foils or concealed on target faces in lineups. Previous research found that with young adults, replication resulted in superior target identification than concealment (Zarkadi et al., 2009). Chapter 4 found the same with young adults but older adults showed similar performance for replication and concealment.

Modelling of the data from Chapter 4 suggested that older adults had a more general representation of faces than young adults, indicating that they were less able to discriminate between similar faces in memory. Older adults did not identify more faces for replication than concealment conditions. Therefore, for replication lineups, the benefit of seeing target faces exactly as presented at study was counteracted by the presence of familiar distinctive features on foil faces. This balance was not the same in young adults where the model parameters indicated a large memory benefit for replication targets.

Chapter 4 is one of the first studies to highlight that age-related associative memory differences need to be considered in terms of their practical implications. The results demonstrated that the nature of older adults' associative memory is qualitatively different to that of young adults, demonstrating that these differences must be considered when testing associative memory in applied contexts.

Preexisting knowledge. In Chapter 5, the specific nature of item and associative memory was explored in terms of their differing reliance on preexisting knowledge. It was hypothesised that item memories are memories involving concepts that already exist in memory (e.g., words) whereas associations are novel links between items in memory. Age deficits in associative memory were therefore defined as age deficits in forming novel connections in the absence of support from preexisting knowledge. Experiment 4 directly tested this hypothesis by comparing item and associative memory in young and older adults where both items and associations were novel and neither could be supported by preexisting knowledge. The novelty of items was manipulated by using nonwords and words as stimuli.

Therefore, for pairs of nonwords, participants would not have encountered the items or their pairings prior to the experiment. This was contrasted to a words condition where participants viewed pairs of unrelated words (similar to many studies in the literature that showed an age-related associative deficit). The words were items that participants had preexisting experience of and their pairings were novel and were not supported by preexisting knowledge.

In the nonwords condition, age deficits were similar for item and associative memory, demonstrating no age-related associative deficit when items and associations are equally novel. For the words condition, age-related associative deficits were present, showing greater age deficits for associative memory than for item memory. The words and the nonwords conditions showed a triple interaction between age, test and condition – the age-related associative deficit was significantly smaller with nonwords than with words.

The results supported the hypothesis that associative deficits in older adults stem from deficits at forming memories in the absence of support from preexisting knowledge. This provided a clear indication of specifically what differs between items and associations in memory, namely their differing reliance on preexisting knowledge. Alternatively, the results could also be explained in terms of dual process accounts of memory. The use of nonwords could have equated young and older adults' use of recollective processes. With nonwords, no meaningful relations can be formed to integrate them in associative memory. Therefore, young adults would not be able to use recollection strategies to aid their memory and this may eliminate their advantage relative to older adults: Dual process accounts postulate that older adults are less able to use recollection (e.g., Yonelinas, 2002) so with nonwords, young adults' associative memory may have been more similar to older adults' associative memory because neither group was using recollection.

Chapter 6 continued the theme of examining preexisting knowledge and its support for memory formation in older adults. The relations between to-be-associated words were manipulated (with and without preexisting knowledge) in a cued recall test of associative memory for links between words. Initially, a new type of relation was considered in a priming study: Integrative relations. These occur when two dissimilar and unassociated words are linked together to form a coherent phrase (e.g., *horse-doctor*). Such relations were crucial to the design as they do not rely on preexisting knowledge in the form of shared features (as semantic relations do), yet they produce clear effects that have previously been shown to facilitate word processing (Estes & Jones, 2009) and memory (Jones et al., 2008) in young adults.

Experiment 5 set out to establish that older adults could use integrative relations to facilitate their responses in a lexical decision task. Older adults completed a priming task where they were primed with a word that was shortly followed by a target word, the task being to decide if the target word was real or a nonword. Integrative and semantic priming were created in trials where the prime and target words were either integratively or semantically related. As a measure of priming, responses to primed trials were compared to baseline trials where the prime was a row of asterisks (i.e., no priming). The results showed semantic priming in line with previous ageing research (Laver & Burke, 1993) and also integrative priming in older adults – semantic and integrative primes led to faster target responses than baseline trials. This was the first demonstration that older adults could rapidly and automatically make use of integrative relations to facilitate responses in a lexical decision task. The results therefore highlighted that older adults may be able to make use of these novel relations to aid associative memory without recourse to preexisting knowledge.

In Experiment 6, young and older adults completed a cued recall task with integrative, semantic, and unrelated word pairs. Participants were shown the pairs in a study phase and later their memory was tested via cued recall where they were shown the left hand word of each pair and asked to recall the right hand word that was originally presented with it. They completed three separate blocks, one with each type of pair relation. The results showed that integrative and semantic word pairs elicited significantly smaller age deficits than unrelated pairs. This demonstrated that relations between words can support associative memory processes in older adults and alleviate age deficits in associative memory. More importantly, unlike semantic word pairs, integrative word pairs form relations that are novel and are unique to the experimental period. Therefore, integrative relations facilitated memory without support from preexisting knowledge.

The results from Chapter 6 initially seem to counter the view from Chapter 5 where preexisting knowledge was considered as responsible for reductions in age deficits. However, when considered in more detail, rather than countering the earlier findings they may actually provide an indication of how preexisting knowledge reduces age differences, namely, by providing effective support to encoding and retrieval processes in older adults. Relations between stimuli may provide a framework upon which to encode associative memory regardless of preexisting knowledge. Additionally, or alternatively, relations may provide support at retrieval: If participants know they are searching for a related word during retrieval, then this would narrow down the material that needed to be searched through. This points towards potentially different roles of environmental support at encoding and retrieval.

Chapter 7 measured the associative deficit in children relative to young adults by testing their item and associative memory. Participants were shown pairs of words and then after a short delay completed item and associative recognition memory tests. Relations between to-be-associated words were also manipulated – semantically related and unrelated word pairs were used as stimuli. Unlike older adults, children only showed small associative memory deficits relative to young adults. Also, manipulating environmental support for associative memory via semantic relations affected children's and young adults' memory similarly. Both groups showed improved associative memory for semantically related word pairs compared to unrelated word pairs. In the literature in general, children often show some form of associative deficit relative to young adults (e.g., Cowan et al., 2006; Sluzenski et al., 2006). Chapter 7 indicated that such a deficit is qualitatively different to that found in older adults in that it is both smaller and less affected by environmental support.

Cognitive resources. Chapter 8 assessed whether available cognitive resources impact item and associative memory performance differently. One of the simplest explanations of the age-related associative deficit is that associative memory requires more cognitive resources than item memory and that older adults' associative memory is poor because they lack cognitive resources (e.g., Troyer et al., 1999).

Experiment 8 attempted to create associative deficits in young adults by reducing the amount of cognitive resources available to them during memory processes. A dual task experiment was set up to divide young adults' attention during encoding and retrieval of item and associative memories. In a memory test using pairs of words as stimuli, a secondary task was conducted during encoding, retrieval or both encoding and retrieval (in three between-subjects conditions). The secondary task was to make semantic judgements about objects and it had previously been shown to disrupt memory when performed alongside encoding and retrieval and to tax prefrontal areas of the brain (areas thought to be responsible for several measures of age-related cognitive decline; see Chapter 1).

When memory performance was compared for full and divided attention, the presence of the secondary task reduced item and associative memory similarly, regardless of whether attention was divided at encoding, retrieval or during both encoding and retrieval. The results indicated that the age-related associative deficit cannot be explained by different levels of cognitive resources required for item and associative memory processes. This is because reducing the amount of cognitive resources available to young adults during memory processes did reduce their memory performance but it did not produce associative deficits as are found in older adults.

Additionally, the results demonstrated that levels of cognitive resources at encoding and retrieval were not differentially important to associative memory. Dual process accounts of age-related associative memory deficits (where age-related recollection deficits are thought to specifically hinder associative memory; see Chapter 2) are based entirely on the retrieval period. The experiment successfully disrupted memory during retrieval but there was no indication that reduced cognitive resources at retrieval leads to greater associative than item memory deficits. This provided indirect evidence against a dual process account of associative deficits.

The key finding from Chapter 8 was that because age-related associative deficits cannot be replicated by a general reduction in cognitive resources, there is likely to be something qualitatively different between age differences in item and associative memory. Ultimately the results provide evidence against a global explanation of age-related associative deficits. Thus there is something specific about associative memory that older adults struggle with.

Environmental Support

A consistent theme throughout the thesis was the notion of environmental support (and internal support from preexisting knowledge) and its impact on age-related associative deficits. Chapter 1 discussed how environmental support minimises age differences in memory tasks and the current results generally found that associative deficits were alleviated by environmental support (cf. Craik, 1982, 1986) and support from preexisting knowledge. The notion of environmental support acting to reduce working memory load was considered in terms of external, stimulus driven support in line with Craik's theory and in terms of internal, knowledge-based support. This is an important point to make as the current research suggests that older adults can make use of preexisting knowledge to reduce working memory load, support memory processes, and reduce their age deficits in memory. Therefore the notion of internal support throughout the thesis is an extension to the original (external support) theory proposed by Craik.

In Chapter 3, older adults were able to show an isolation effect (where isolation was based on associative memory) that was similar to young adults, possibly because the distinctiveness of isolates at encoding provided environmental support to older adults, therefore minimising age differences. In Chapter 5, when item and associative tests were equated on preexisting knowledge by using novel associations *and* novel items, age-related associative deficits were not present. This could be because the item and associative memory were equated on the basis of support from preexisting knowledge. Chapter 6 demonstrated that when to-be-associated items were easily related to each other, age deficits in associative memory were reduced. This was possibly because the relations provided effective support to the encoding and retrieval of associations. Overall, this evidence points towards a distinction between item and associative memory, namely, that associative memory is less likely than item memory to be supported by environmental factors and preexisting knowledge. When item and associative tests are equated in terms of support available to memory processes, age-related associative memory deficits are less apparent.

Some parts of the thesis indicated how environmental support differed from other explanations of age-related associative deficits. Chapter 8 showed that reducing cognitive resources disrupted all memory in young adults, not just associative memory. The results demonstrated that the age-related associative deficit cannot be explained in terms of different levels of cognitive resources available to young and older adults. This provided evidence against the possibility that environmental support alleviates associative memory deficits in older adults simply by reducing the amount of cognitive resources required to form associative memory. Also Chapter 7 indicated that memory deficits in children are qualitatively different to those found in older adults: Manipulation of environmental support affected children's and young adults' associative memory similarly. Therefore the benefit of environmental support to older adults' associative memory may be specific to problems caused by an age-related associative memory deficit.

Neuropsychological Deficits

Chapter 2 discussed how degradation of prefrontal and medial temporal/hippocampal areas in old age may lead to associative deficits. The results from Chapter 8 indicate that prefrontal deficits may not be sufficient to produce larger associative deficits than item deficits in older adults. Item and associative memory were disrupted in young adults with a concurrent task. The task used to disrupt memory required semantic judgements of whether an object was a living or non-living thing and such judgements have been shown to activate prefrontal areas (e.g., Kapur et al., 1994). The task did not disrupt associative memory more than item memory which suggests that prefrontal disruption does not differentially affect the two memory types. Therefore, age-related associative deficits may be affected more by medial temporal/hippocampal degradation, although the current results only provide indirect evidence for such a conclusion. Also, contrary to this view is the finding that patients with prefrontal lesions show relational memory deficits (Cabeza, 2006).

The research into associative deficits in children (Chapter 7) also indicated that an age-related associative deficit may be more dependent on medial temporal/hippocampal functionality. In Experiment 7 children did not show associative deficits relative to young adults to the same extent that older adults do and their associative memory was less affected by environmental support than has been shown in older adults. Memory deficits in children have been considered to result from protracted strategic/prefrontal development rather than protracted medial temporal/hippocampal development (Casey et al., 2005; Shing et al., 2010). Again this provides some evidence towards the idea that the age-related associative deficit may be more dependent on medial temporal/hippocampal functionality.

Speed

The majority of studies reported in this thesis also collected a measure of information processing speed (Digit Symbol Substitution task, DSST, Wechsler, 1981) to gain an overview of individual differences between young and older participants. As expected, there were always highly significant differences between young and older age groups in terms of their DSST scores. Importantly, processing speed has been shown to account for a large proportion of age-related variance in many cognitive tasks including working memory (e.g., Salthouse, 1996; see Chapter 1). However, this was generally not the case for measures of associative memory in the present studies. To take one illustrative example, in Experiment 6 there was a significant correlation between participants' exact ages and their cued recall scores for unrelated word pairs, $r(72) = -.68$, $p < .001$. With DSST scores partialled out, the correlation remained highly significant, $r = -.43$, $df = 69$, $p < .001$, indicating age-related deficits in associative memory above and beyond general slowing.

Global Deficits Versus Associative Deficits

In Chapter 8 it was shown that a general reduction of cognitive resources did not impact associative memory more than item memory in young adults. Despite this and many similar findings in the literature (see Chapter 8), it has been argued that the associative deficit in older adults may simply be because the nature of associative memory tests makes them more sensitive to general deficits in cognition. Benjamin (2010) has challenged empirical dissociations between associative/item memory and context/content memory performance in young and older adults. He argues that item and associative memory are part of a continuum and describes a '[r]epresentational sparsity hypothesis: Stimuli, situations, or events that are less central to the rememberer's tasks, goals, and perceptual and attentional biases are represented more sparsely in memory' (Benjamin, 2010, p. 1062). With this hypothesis, item and associative memory are seen as one and the same thing varying in their degree of representational sparsity – associative memories are considered more sparse and hence more difficult for older adults to remember as well as young adults.

Benjamin (2010) postulates that general global deficits in older adults' memory are exaggerated for associative memories because of their greater sparsity, the level of detail used to store a given piece of information. It was hypothesised that associative memories are generally less focal to attention and ongoing behaviour than item memories and that they are therefore represented in less detail. By modelling this hypothesis, Benjamin showed that memory representations that are more sparse are more sensitive to disruptions of overall memory fidelity. Benjamin argued that this showed how a global reduction in fidelity due to cognitive ageing can account for specific age-related associative deficits in empirical research.

In a follow up paper, Benjamin, Diaz, Matzen, and Johnson (2011) presented some behavioural evidence to show that contextual information is represented more sparsely than content information. They argued that when encoding occurs on a shorter time scale, concepts are encoded with a lower fidelity to the greater detriment of sparse/context memory. Results with young adults showed that reducing study time impacted context memory performance more than item memory performance. However, such a result is incongruent with that of Chapter 8 and many other divided attention studies (see Chapter 8 for review), which generally show that divided attention reduces item and associative memory similarly. If time available for encoding impacts sparse memory more than richer memory then the same should be true for dividing attention at encoding but it is not. Furthermore, it was noted above that processing speed does not fully account for age-related associative memory deficits.

Benjamin et al. (2011) also showed that age differences are exaggerated for information that is not the focus of attention. This is perhaps in line with the isolation results from Chapter 3 where older adults demonstrated an ability to memorise isolates that gained their attention. But again this is incongruent with prior research: Naveh-Benjamin (2000) found that age-related associative deficits were *reduced* when the associative test was incidental compared to explicit in the memory instructions. Other studies have shown similar results (e.g., Old & Naveh-Benjamin, 2008a). The representational sparsity hypothesis introduces a new problem in that it requires us to define how different types of memory stimuli are represented in memory with different levels of sparsity, which is a difficult and abstract problem to solve.

Speculation and Future Directions

The most striking results from the thesis come from investigations into the differences between item and associative memory. Similar to Benjamin and colleagues (Benjamin, 2010; Benjamin et al., 2011), the current results indicate that item and associative memory are represented differently. However, rather than the two memory types differing in their level of sparsity, the current results indicate that item and associative memory are subject to different levels of environmental support during memory processing. Defining item and associative memory as different in structure (rather than as relying on different processes or as stored in different locations) is appealing in that it allows age-related associative deficits to be explained in terms of global accounts of age-related memory decline such as deficits in self-initiated processing (Craik, 1986), and inhibitory deficits (Hasher & Zacks, 1988). This is because the way item and associative information are physically represented in memory may cause the two memory types to be treated differently by a single, global memory process. Therefore, even if the fundamental rules of processing are the same for both types of memory, they may appear to be dissociated behaviourally due to different implementation of processes for different memory structures.

An important result of Naveh-Benjamin (2000) was to extend the finding of age deficits in context memory (e.g., Spencer & Raz, 1995) to age deficits in associative memory. This demonstrated that contextual memory deficits could be a result of more basic associative memory deficits, therefore highlighting that contextual information is not unique or special in terms of cognitive decline. By showing age-related deficits in the binding of arbitrary items, Naveh-Benjamin's (2000) results showed that it is not the type of information that yields age deficits but rather the necessity to bind units of information together. Initially, such a result seems to point towards associative memory as a fundamentally different process to item memory, contrary to the view expressed above. However, the current results point to a differential reliance on environmental support and support from preexisting knowledge for item and associative memory. It is hypothesised here that this difference affects the way item and associative memory are processed, leading to age-related associative deficits.

The predictability of retrieval processes. It is postulated here that the age-related associative deficit arises from differences in item and associative memory structure. It is hypothesised that cognitive processes take into account a relationship between encoding and retrieval in order to minimise the amount of information that needs to be stored in memory. This in turn allows an individual to minimise the amount of memory resources required to store and retrieve information. In relation to an age-related associative deficit, the hypothesis indicates that concepts supported by preexisting knowledge or environmental factors require fewer new connections in memory. Thus item memory involves fewer new connections than associative memory and is therefore less susceptible to global memory decline as a result of cognitive ageing.

The hypothesis is consistent with the encoding specificity principle (Tulving & Thompson, 1973), namely that '[r]emembering depends on the interaction between the conditions at encoding and the conditions at retrieval'(Neath & Surprenant, 2005, p. 223). It is suggested here that when the nature of retrieval processes can be predicted at encoding, memory processes can be more efficient. For example, if you wanted to remember to buy a family member a birthday card you might write a note that simply says 'b-card' and place it somewhere that it will be encountered later. When you read the note later, you will then think about the intended recipient of the card, remember the date of their birthday, plan a journey to an appropriate shop and so on. All of these concepts were predictably activated at retrieval by a simple cue 'b-card'. It is assumed that if such a piece of information was encoded in memory rather than put on a note that the information encoded would be similar to the content of the note (i.e., a minimal, efficient cue). The point is that if the retrieval process can be relied upon to interpret and act appropriately on a minimal cue, then the encoding process needs only to encode that minimal information. Therefore, the closer that the behaviour at retrieval can be predicted at the time of encoding, the more effective and efficient the encoding process can be.

The hypothesis suggests that the more stable and predictable a retrieval process is, the easier it is to encode a cue that will result in successful retrieval. This is in line with the notion of environmental support enhancing memory, a factor that also alleviates age deficits in memory: Environmental support (and preexisting knowledge) guides encoding and retrieval processes, which would make them more stable and predictable. In terms of item and associative memory, it is hypothesised that item memory retrieval is more predictable than associative memory retrieval. Typical measures of associative deficits show young and older adults memory sets consisting of pairs of unrelated words for later recognition tests of item memory (old vs. new words) and associative memory (intact vs. recombined pairs of words). To remember an item, a participant must remember the word and its presentation during the experiment – this requires a temporal recency modification to the preexisting item (word) in memory. In the corresponding item recognition test, the participant must evaluate each old/new item and search for a sense of temporal recency. This is a very specific and predictable search, which would hypothetically be cued by minimal modifications to memory during encoding.

To remember the association between two unrelated words the participant must modify at least one of the words in preexisting memory so that it evokes the other during retrieval. For the associative recognition test, when presented with a pair of words in order to recognise intact/recombined associations, retrieval cannot search simply for a sense of temporal recency as all individual stimuli are old. Successful associative memory retrieval requires memory for an entirely new concept that incorporates the two words accurately enough to distinguish them from recombined word pairs. Such a concept could take many forms; it may be a concept for a mediator that relates to the words, a sentence that contains the words or it may be a concept for some sort of imagery, story or event that links the words and so on. The retrieval process for associations has a lot more to search through so to cover all of the possibilities more new memories must be created at encoding to guide retrieval. That is, during encoding, there is less information about how retrieval will search for an association. Thus to improve the chance of retrieval encountering an encoded event, encoding must make more modifications. Therefore, in line with research showing larger age deficits for more complex tasks (e.g., Salthouse, 1988), age differences may be greater for associative memory than item memory because overall the memory demand is higher for associative memory (see Naveh-Benjamin, 2000). The novel aspect of the current hypothesis is that it defines why associative memory is more complex – because associative retrieval processes are less predictable at the time of encoding than item retrieval processes.

The hypothesis also can be used to explain the results from Chapter 6 and other similar studies that manipulate relations between to-be-associated stimuli. In Chapter 6, relations between words made associative memory easier and reduced the age-related associative deficit with cued recall. The presence of relations between stimuli allows retrieval processes to become more predictable. The retrieval process will search for a word that is related to the cue and this allows the encoding process to specifically create a cue that appeals to such a search (i.e., by encoding the cue, the target and their relation). The current hypothesis also predicts that the awareness of a relation is sufficient to increase the predictability of retrieval. This explains how integrative and semantic relations aid associative memory similarly, despite being based on different levels of preexisting knowledge.

Ultimately, the current hypothesis indicates that when retrieval processes are more predictable at the time of encoding, memory will improve in general and age differences will be reduced. In terms of relations between to-be-associated items, the hypothesis would predict that if relations were not certain to be present (for example in a mixed list containing unrelated and semantically related word pairs), the behaviour of the retrieval process would be less predictable during encoding and the benefit of relations would be smaller than in a list that entirely contains related word pairs. Therefore, in a mixed list, the reduction in age-related associative deficits with related words would be less than that seen in Chapter 6 where blocked sets of related word pairs aided associative memory. Also for recognition tests of associative memory, the hypothesis predicts that if the recognition of associations could work in a more structured way, memory would improve and age deficits would reduce. For example, if participants were instructed to always encode associations between two words with a mediator, this would narrow down the possible behaviour at retrieval (retrieval would search specifically for a mediator) and allow encoding to be more accurate/efficient, resulting in a smaller age deficit. As has been discussed in Chapter 6, Naveh-Benjamin et al. (2007) reduced the age-related associative deficit by encouraging strategy use at encoding and retrieval.

The hypothesis is also able to make predictions that existing descriptions of memory processes could not account for. For example, an informative test of the hypothesis for future research would be to evaluate a counter-intuitive prediction that it makes. Consider the following list of word pairs: *car-spoon*, *table-tree*, *blanket-hat*, *pen-keyboard*. Initially these pairings may seem unrelated but both objects in each pair are largely made of the same material (metal, wood, fabric, and plastic, respectively). There is a relation between the words but it is not apparent at the time of encoding. The hypothesis suggests that if these relations are not apparent at encoding then they will have no benefit to memory processes compared to unrelated word pairs. This is because the encoding process will not be able to predict a search for related items at retrieval (i.e., the relations do not increase the predictability of retrieval processes). The counter-intuitive prediction is that if participants were informed about the nature of the relations before retrieval (but after encoding) there would *still* be no benefit of relatedness because it was not taken into account at encoding. Therefore, in an experiment, if one group of participants was informed about the relations before encoding, they would show a benefit of relatedness but another group who only learned the nature of the relations immediately before retrieval would show no benefit. Additionally, the hypothesis predicts that an age-related associative deficit would not be reduced by relations if they were only brought to awareness after encoding.

Future research. The logical goal of research into the age-related associative deficit should be to link it with other theories of cognitive ageing. It is likely that some form of global decline has a greater impact on associative than item memory. The impact may be upon the way memories are processed or it may be upon the way memories are represented/stored. There is much research in the literature to support a specific deficit in associative memory (Old & Naveh-Benjamin, 2008a). However, if item and associative information is represented differently in memory, associative memory tests may be more sensitive to global age-related memory decline than item memory tests (Benjamin, 2010; Benjamin et al., 2011). That is, age-related associative memory deficits may be a result of the experimental paradigms commonly used to assess memory performance. Until it is proved otherwise, a more parsimonious global explanation of age-related memory decline should be pursued.

The results from this thesis suggest that defining the specific differences in the way item and associative memories are represented and processed is a valid direction for research into the age-related associative deficit. Considering the predictability of retrieval processes at the time of encoding may yield new insights into the interaction between encoding and retrieval and align our understanding of age-related global and associative deficits.

References

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: 1. Evidence from divided attention costs. *Psychology and Aging, 13*, 405-423.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189-208.

Baddeley, A. D., Grant, S., Wight, E., & Thompson, N. (1973). Imagery and visual working memory. In P. M. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 205-217). London: Academic Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. Bower (Ed.), *Recent advances in learning and motivation* (Vol. VIII, pp. 47-90). New York: Academic Press.

Baddeley, A. D., Vivien, L., Eldridge, M., & Thompson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113*, 518-540.

Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D., & Roediger III, H. L. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging, 21*, 19-31.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445-459.

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, *118*, 175-192.

Bartlett, J. C., Leslie, J. E., Tubbs, A., & Fulton, A. (1989). Aging and memory for pictures of faces. *Psychology and Aging, 4*, 276-283.

Bastin, C., & Van der Linden, M. (2005). Memory for temporal context: Effects of ageing, encoding instructions, and retrieval strategies. *Memory, 13*, 95-109.

Bastin, C., & Van der Linden, M. (2006). The effects of aging on the recognition of different types of associations. *Experimental Aging Research, 32*, 61-77.

Bender, A. R., Naveh-Benjamin, M., & Raz, N. (2010). Associative deficit in recognition memory in a lifespan sample of healthy adults. *Psychology and Aging*, *25*, 940-948.

Benjamin, A. S. (2010). Representational explanations of “process” dissociations in recognition: The DRYAD theory of aging and memory judgments. *Psychological Review, 117*, 1055-1079.

Benjamin, A. S., & Craik, F. I. M. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition*, *29*, 691-697.

Benjamin, A. S., Diaz, M., Matzen, L. E., & Johnson, B. (2011, August 29). Tests of the DRYAD theory of the age-related deficit in memory for context: Not about context, and not about aging. *Psychology and Aging*, Advance online publication. doi: 10.1037/a0024786.

Bireta, T. J., Surprenant, A. M., & Neath, I. (2008). Age-related differences in the von Restorff isolation effect. *Quarterly Journal of Experimental Psychology, 61*, 345-352.

Birren, J. E. (1965). Age changes in speed of behaviour: Its central nature and physiological correlates. In A. T. Welford & J. E. Birren (Eds.), *Behaviour,* *aging, and the nervous system* (pp. 191-216). Springfield, IL: Charles C Thomas.

BNC. (2007). The British National Corpus, version 3 (BNC XML Edition) [Data set]. Retrieved from http://www.natcorp.ox.ac.uk/.

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW). Gainesville, FL. The NIMH Center for the Study of Emotion and Attention, University of Florida.

Braver, T. S., & West, R. (2008). Working memory, executive control, and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 311-372). New York: Psychology Press.

Brockmole, J. R., Parra, M. A., Della Sala, S., & Logie, R. H. (2008). Do binding deficits account for age-related decline in visual working memory? *Psychonomic Bulletin & Review, 15*, 543-547.

Burke, D. M. (1997). Language, aging, and inhibitory deficits: Evaluation of a theory. *Journal of Gerontology: Psychological Sciences, 52B*, P254-P264.

Burke, D. M., MacKay, D. G., & James, L. E. (2000). Theoretical approaches to language and aging. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 204-237). New York: Oxford University Press.

Cabeza, R. (2006). Prefrontal and medial temporal lobe contributions to relational memory in young and older adults. In H. D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Handbook of binding and memory: Perspectives from cognitive neuroscience* (pp. 595-626). New York: Oxford University Press.

Cabeza, R., Anderson, N. D., Houle, S., Mangels, J. A., & Nyberg, L. (2000). Age-related differences in neural activity during item and temporal-order memory retrieval: A positron emission tomography study. *The Journal of Neuroscience, 12*, 197-206.

Cabeza, R., Grady, C. L., Nyberg, L., McIntosh, A. R., Tulving, E., Kapur, S., et al. (1997). Age-related differences in neural activity during memory encoding and retrieval: A positron emission tomography study. *The Journal of Neuroscience, 17*, 391-400.

Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition - A maturational component? *Developmental Psychology, 16*, 257-269.

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: What have we learned about cognitive development? *Trends in Cognitive Sciences, 9*, 104-110.

Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging, 20*, 718-721.

Castel, A. D. (2007). Aging and memory for numerical information: The role of specificity and expertise in associative memory. *Journal of Gerontology: Psychological Sciences, 62B*, P194-P196.

Castel, A. D., & Craik, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging, 18*, 873-885.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin, 98*, 67-83.

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24*, 403-416.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55-81.

Chung, M. S., & Thomson, D. M. (1995). Development of face recognition. *British Journal of Psychology, 86,* 55-87.

Cimbalo, R. S., & Brink, L. (1982). Aging and the von Restorff isolation effect in short/term memory. *Journal of General Psychology, 106,* 69-76.

Cohen, G. (1990). Recognition and retrieval of proper names: Age differences in the fan effect. *European Journal of Cognitive Psychology, 2,* 193-204.

Cohn, M., Emrich, S. M., & Moscovitch, M. (2008). Age-related deficits in associative memory: The influence of impaired strategic retrieval. *Psychology and Aging, 23,* 93-103.

Conrad, J. (1964a). Acoustic confusion in immediate memory. *British Journal of Psychology, 55,* 75-84.

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

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology, 42,* 1089-1102.

Craik, F. I. M. (1982). Selective changes in encoding as a function of reduced processing capacity. In F. Klix, J. Hoffman & E. van der Meer (Eds.), *Cognitive research in psychology* (pp. 152–161). Berlin: DVW.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities, mechanisms, and performance* (pp. 409-422). Amsterdam: Elsevier.

Craik, F. I. M. (2000). Age-related changes in human memory. In D. C. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 75-92). Philadelphia, PA: Psychology Press.

Craik, F. I. M. (2006). Remembering items and their contexts: Effects of ageing and divided attention. In H. D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Handbook of binding and memory: Perspectives from cognitive neuroscience* (pp. 571-594). New York: Oxford University Press.

Craik, F. I. M., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences, 10*, 131-138.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 191-211). New York: Plenum Press.

Craik, F. I. M., Luo, L., & Sakuta, Y. (2010). Effects of aging and divided attention on memory for items and their contexts. *Psychology and Aging, 25*, 968-979.

Craik, F. I. M., & McDowd, J. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 474-479.

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

Dennis, N. A., & Cabeza, R. (2008). Neuroimaging of healthy cognitive aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 1-54). New York: Psychology Press.

Detterman, D. K. (1975). The von Restorff effect and induced amnesia: Production by manipulation of sound intensity. *Journal of Experimental Psychology: Human Learning and Memory, 1*, 614-628.

Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging, 13*, 597-607.

Dunlosky, J., & Hertzog, C. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition, 29,* 247-253.

Dunlosky, J. R., Hunt, R. R., & Clark, E. (2000). Is perceptual salience needed in explanations of the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 649-657.

Erickson, R. L. (1963). Relational isolation as a means of producing the von Restorff effect in paired-associate learning. *Journal of Experimental Psychology, 66,* 111-119.

Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic thinking: The apprehension and consequences of thematic relations. In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 54, pp. 249-294). Burlington: Academic Press.

Estes, Z., & Jones, L. L. (2009). Integrative priming occurs rapidly and uncontrollably during lexical processing. *Journal of Experimental Psychology: General*, *138*, 112-130.

Fastenau, P. S., Denburg, N. L., & Abeles, N. (1996). Age differences in retrieval: Further support for the resource-reduction hypothesis. *Psychology and Aging, 11*, 140-146.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fernandes, M. A., & Moscovitch, M. (2000). Divided attention and memory: evidence of substantial interference effects at retrieval and encoding. *Journal of Experimental Psychology: General, 129,* 155-176.

Fernandes, M. A., & Moscovitch, M. (2002). Factors modulating the effect of divided attention during retrieval of words. *Memory & Cognition, 30,* 731-744.

Fernandes, M. A., & Moscovitch, M. (2003). Interference effects from divided attention during retrieval in younger and older adults. *Psychology and Aging, 18,* 219-230.

Fernandes, M. A., Moscovitch, M., Ziegler, M., & Grady, C. (2005). Brain regions associated with successful and unsuccessful retrieval of verbal episodic memory as revealed by divided attention. *Neuropsychologia, 43,* 1115-1127.

Fernandez, G., & Fell, J. (2006). Rhinal-hippocampal contribution to declarative memory formation. In H. D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Handbook of binding and memory: Perspectives from cognitive neuroscience* (3rd ed., pp. 163-190). New York: Oxford University Press.

Fisher, D. L., Duffy, S. A., & Katsikopoulos, K. V. (2000). Cognitive slowing among older adults: What kind and how much? In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive Aging* (pp. 87-122). New York: Oxford University Press.

Flin, R. H. (1980). Age effects for children's memory for unfamiliar faces. *Developmental Psychology, 16,* 373-374.

Flin, R. H. (1985). Development of visual memory: An early adolescent regression. *The Journal of Early Adolescence*, *5*, 259-266.

Garden, S. E., Phillips, L. H., & MacPherson, S. E. (2001). Midlife aging, open-ended planning, and laboratory measures of executive function. *Neuropsychology, 15,* 472-482.

Geraci, L., McDaniel, M. A., Manzano, I., & Roediger III, H. L. (2009). The influence of age on memory for distinct events. *Memory & Cognition, 37,* 175-180.

Geraci, L., & Rajaram, S. (2002). The orthographic distinctiveness effect on direct and indirect tests of memory: Delineating the awareness and processing requirements. *Journal of Memory and Language, 47,* 273-291.

Gerard, L., Zacks, R. T., Hasher, L., & Radvansky, G. A. (1991). Age deficits in retrieval: The fan effect. *Journal of Gerontology: Psychological Sciences, 46,* P131-P136.

Gibson, E. J. (1940). A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychological Review, 47*, 196-229.

Gilbert, J. G. (1941). Memory loss in senescence. *The Journal of Abnormal and Social Psychology, 36*, 73-86.

Glisky, E. L., & Kong, L. L. (2008). Do young and older adults rely on different processes in source memory tasks? A neuropsychological study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 809-822.

Glisky, E. L., Rubin, S. R., & Davidson, P. S. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1131-1146.

Glucksberg, S., & Estes, Z. (2000). Feature accessibility in conceptual combination: Effects of context-induced relevance. *Psychonomic Bulletin & Review, 7,* 510-515.

Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory, 6,* 225-255.

Grady, C. L., McIntosh, A. R., Horwitz, B., Maisog, J. M., Ungerleider, L. G., Mentis, M. J., … Haxby, J.V. (1995). Age-related reductions in human recognition memory due to impaired encoding. *Science, 269*, 218-221.

Green, R. T. (1956). Surprise as a factor in the von Restorff effect. *Journal of Experimental Psychology, 52,* 340-344.

Green, R. T. (1958). Surprise, isolation, and structural change as factors affecting recall of a temporal series. *British Journal of Psychology, 49*, 21-30.

Guerard, K., Hughes, R. W., & Tremblay, S. (2008). An isolation effect in serial memory for spatial information. *The Quarterly Journal of Experimental Psychology, 61*, 752-762.

Gumenik, W. E., & Levitt, J. (1968). The von Restorff effect as a function of difference of the isolated item. *American Journal of Psychology, 81*, 247-252.

Gunning-Dixon, F. M., & Raz, N. (2000). The cognitive correlates of white matter abnormalities in normal aging: A quantitative review. *Neuropsychology, 14,* 224-232.

Gutchess, A. H., & Park, D. C. (2009). Effects of ageing on associative memory for related and unrelated pictures. *European Journal of Cognitive Psychology, 21*, 235-254.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 163-169.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356-388.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193-225). New York: Academic Press.

Healy, M. R., Light, L. L., & Chung, C. (2005). Dual-process models of associative recognition in young and older adults: Evidence from receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 768-788.

Hertzog, C., & Hultsch, D. F. (2000). Metacognition in adulthood and old age. In T. A. Salthouse & F. I. M. Craik (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 417-466). Mahwah, NJ: Erlbaum Associates, Inc.

Hess, T. M., & Slaughter, S. J. (1990). Schematic knowledge influences on memory for scene information in young and older adults. *Developmental Psychology, 26*, 855-865.

Hockley, W. E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory & Cognition, 27*, 657-664.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107-129.

Horne, J. A., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*, 97-110.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging, 5*, 356-368.

Humphreys, M. S. (1976). Relational information and the context effect in recognition memory. *Memory & Cognition, 4*, 221-232.

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review, 2*, 105-112.

Hunt, R. R., & Lamb, C. A. (2001). What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1359-1366.

Iidaka, T., Sadato, N., Yamada, H., Murata, T., Omori, M., & Yonekura, Y. (2001). An fMRI study of the functional neuroanatomy of picture encoding in younger and older adults. *Cognitive Brain Research, 11*, 1-11.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513-541.

Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 3-22.

James, L. E., Fogler, K. A., & Tauber, S. K. (2008). Recognition memory measures yield disproportionate effects of aging on learning face-name associations. *Psychology and Aging*, *23*, 657-664.

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging, 8,* 283-293.

Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience*, *2*, 475-483.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28.

Jones, F. N., & Jones, M. H. (1942). Vividness as a factor in learning lists of nonsense syllables. *The American Journal of Psychology, 55*, 96-101.

Jones, L. L., Estes, Z., & Marsh, R. L. (2008). An asymmetric effect of relational integration on recognition memory. *The Quarterly Journal of Experimental Psychology, 61*, 1169-1176.

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

Kapur, S., Craik, F. I., Tulving, E., Wilson, A. A., Houle, S., & Brown, G. M. (1994). Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect. *Proceedings of the National Academy of Sciences, 91*, 2008-2011.

Kelley, M. R., & Nairne, J. S. (2001). von Restorff revisited: Isolation, generation, and memory for order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 54-66.

Kersten, A. W., & Earles, J. L. (2010). Effects of aging, distraction, and response pressure on the binding of actors and actions. *Psychology and Aging, 25*, 620-630.

Kilb, A., & Naveh-Benjamin, M. (2007). Paying attention to binding: Further studies assessing the role of reduced attentional resources in the associative deficit of older adults. *Memory & Cognition, 35*, 1162-1174.

Kimble, G. A., & Dufort, R. H. (1955). Meaningfulness and isolation as factors in verbal learning. *Journal of Experimental Psychology, 50*, 361-368.

Knapp, B. R., Nosofsky, R. M., & Busey, T. A. (2006). Recognizing distinctive faces: A hybrid-similarity exemplar model account. *Memory & Cognition, 34*, 877-889.

Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging, 25*, 498-503.

Kramer, A. F., Humphrey, D. G., Larish, J. F., Logan, G. D., & Strayer, D. L. (1994). Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. *Psychology and Aging, 9*, 491-512.

Kramer, A. F., Larish, J. F., & Strayer, D. L. (1995). Training for attentional control in dual task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied, 1*, 50-76.

Kramer, A. F., Larish, J. F., Weber, T. A., & Bardell, L. (1999). Training for executive control: Task coordination strategies and aging. In D. Gopher & A. Koriat (Eds.), *Attention and Performance: XVIII* (pp. 617-652). Cambridge, MA: The MIT Press.

Kramer, A. F., & Madden, D. J. (2008). Attention. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of ageing and cognition* (3rd ed., pp. 189-249). New York: Psychology Press.

Lamberts, K. (2005). Mathematical modelling of cognition. In K. Lamberts & R. L. Goldstone (Eds.), *The handbook of cognition* (pp. 407-421). London: SAGE Publications Ltd.

Laver, G. D. (2009). Adult aging effects on semantic and episodic priming in word recognition. *Psychology and Aging, 24*, 28-39.

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging, 8*, 34-43.

Lewandowsky, S., Nimmo, L., & Brown, G. D. A. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language, 58*, 415-428.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42*, 333-376.

Light, L. L., Patterson, M. M., Chung, C., & Healy, M. R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory & Cognition, 32*, 1182-1193.

Light, L. L., Prull, M. W., La Voie, D. J., & Healy, M. R. (2000). Dual-process theories of memory in old age. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 238-300). New York: Oxford University Press.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9*, 339-355.

Lindenberger, U., & Baltes, P. B. (1997). Intellectual functioning in old and very old age: Cross-sectional results from the Berlin Aging Study. *Psychology and Aging, 12*, 410-432.

Lindenberger, U., Marsiske, M., & Baltes, P. B. (2000). Memorizing while walking: Increase in dual-task costs from young adulthood to old age. *Psychology and Aging, 15*, 417-436.

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology, 52*, 297-318.

Lloyd, M. E., Doydum, A. O., & Newcombe, N. S. (2009). Memory binding in early childhood: Evidence for a retrieval deficit. *Child Development, 80*, 1321-1328.

Logan, J. M., & Balota, D. A. (2008). Expanded vs. equal interval spaced retrieval practice: Exploring different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology, and Cognition, 15*, 257-280.

Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C., & Buckner, R. L. (2002). Under-recruitment and nonselective recruitment: Dissociable neural mechanisms associated with aging. *Neuron, 33*, 827-840.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*, 203-208.

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.

Luszcz, M. A., Roberts, T. H., & Mattiske, J. (1990). Use of relational and item-specific information in remembering by older and younger adults. *Psychology and Aging, 5*, 242-249.

MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: A theory of new learning and the use of old connections. In T. M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 213-263). Amsterdam: Elsevier.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98*, 185-199.

Mann, V. A., Diamond, R., & Carey, S. (1979). Development of voice recognition: Parallels with face recognition. *Journal of Experimental Child Psychology, 27*, 153-165.

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology, 16*, 437-458.

May, C. P., Kane, M. J., & Hasher, L. (1995). Determinants of negative priming. *Psychological Bulletin, 118*, 35-54.

Maylor, E. A., Watson, D. G., & Hartley, E. L. (2011). Effects of distraction on visual enumeration in children and adults. *Developmental Psychology, 47*, 1440-1447.

McDowd, J. M. (1997). Inhibition in attention and aging. *Journal of Gerontology: Psychological Sciences, 52B*, P265-P273.

McDowd, J. M., & Oseas-Kreger, D. M. (1991). Aging, inhibitory processes, and negative priming. *Journal of Gerontology: Psychological Sciences, 46*, P340-P345.

McDowd, J. M., & Shaw, R. J. (2000). Attention and aging: A functional perspective. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 221-292). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Memon, A., & Bartlett, J. (2002). The effects of verbalization on face recognition in young and older adults. *Applied Cognitive Psychology, 16*, 635-650.

Memon, A., Gabbert, F., & Hope, L. (2004). The ageing eyewitness. In J. R. Adler (Ed.), *Forensic Psychology: Debates, Concepts and Practice* (pp. 96-112). Uffcolme, Devon: Willan, Forensic Psychology Series.

Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: The effects of mugshot viewing and choosing in young and old adults. *Memory & Cognition, 30*, 1219-1227.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Mitchell, K. J., Johnson, M. K., Raye, C. L., & D'Esposito, M. (2000). fMRI evidence of age-related hippocampal dysfunction in feature binding in working memory. *Cognitive Brain Research, 10*, 197-206.

Mitchell, K. J., Johnson, M. K., Raye, C. L., Mather, M., & D'Esposito, M. (2000). Aging and reflective processes of working memory: Binding and test load deficits. *Psychology and Aging, 15*, 527-541.

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience, 4*, 257-267.

Mulligan, N. W., & Osborn, K. (2009). The modality-match effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 564-571.

Murphy, K. R., & Myors, B. (1998). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests.* Mahwah, NJ: Lawrence Erlbaum.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679-684.

Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging, 7*, 257-270.

Nachmias, J., Gleitman, H., & McKenna, V. V. (1961). The effect of isolation of stimuli and responses in paired associates. *American Journal of Psychology, 74*, 452-456.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1170-1187.

Naveh-Benjamin, M. (2006). Binding of memories: Adult-age differences and the effect of divided attention in young adults on episodic memory. In H. D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Handbook of binding and memory: Perspectives from cognitive neuroscience* (pp. 627-656). New York: Oxford University Press.

Naveh-Benjamin, M., Brav, T. K., & Levy, O. (2007). The associative memory deficit of older adults: The role of strategy utilization. *Psychology and Aging, 22*, 202-208.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Kreuger, S. (2005). Divided attention in younger and older adults: Effects of strategy and relatedness on memory performance and secondary task costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 520-537.

Naveh-Benjamin, M., Guez, J., Kilb, A., & Reedy, S. (2004). The associative memory deficit of older adults: Further support using face–name associations. *Psychology and Aging, 19*, 541-546.

Naveh-Benjamin, M., Guez, J., & Marom, M. (2003). The effects of divided attention at encoding on item and associative memory. *Memory & Cognition, 31*, 1021-1035.

Naveh-Benjamin, M., Guez, J., & Shulman, S. (2004). Older adults' associative deficit in episodic memory: Assessing the role of decline in attentional resources. *Psychonomic Bulletin & Review, 11*, 1067-1073.

Naveh-Benjamin, M., Hussain, Z., Guez, J., & Bar-On, M. (2003). Adult age differences in episodic memory: Further support for an associative-deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 826-837.

Naveh-Benjamin, M., Shing, Y. L., Kilb, A., Werkle-Bergner, M., Lindenberger, U., & Li, S.-C. (2009). Adult age differences in memory for name-face associations: The effects of intentional and incidental learning. *Memory, 17,* 220-232.

Neath, I., & Surprenant, A. M. (2005). *Mechanisms of memory*. London: SAGE Publications Ltd.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*, 402-407.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*, 838-848.

Nosofsky, R. M., & Zaki, S. R. (2003). A hybrid-similarity exemplar model for predicting distinctiveness effects in perceptual old-new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1194-1209.

O'Keefe, D. J. (2007). Post hoc power, observed power, a priori power, retrospective power, prospective power, achieved power: Sorting out appropriate uses of statistical power analysis. *Communication Methods and Measures, 1*, 291-299.

Oeppen, J., & Vaupel, J. W. (2002). Broken limits to life expectancy. *Science, 296*, 1029-1031.

Old, S. R., & Naveh-Benjamin, M. (2008a). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*, 104-118.

Old, S. R., & Naveh-Benjamin, M. (2008b). Memory for people and their actions: Further evidence for an age-related associative deficit. *Psychology and Aging, 23*, 467-472.

Overman, A. A., & Becker, J. T. (2009). The associative deficit in older adult memory: Recognition of pairs is not improved by repetition. *Psychology and Aging, 24*, 501-506.

Park, D. C. (2000). The basic mechanisms accounting for age-related decline in cognitive function. In D. C. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 3-21). Philadelphia, PA: Psychology Press.

Park, D. C., Smith, A. D., Dudley, W. N., & Lafronza, V. N. (1989). Effects of age and a divided attention task presented during encoding and retrieval on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1185-1191.

Park, D. C., Smith, A. D., Morrell, R. W., Puglisi, J. T., & Dudley, W. N. (1990). Effects of contextual integration on recall of pictures by older adults. *Journal of Gerontology: Psychological Sciences, 45*, P52-P57.

Parkin, A. J., & Java, R. I. (2000). Determinants of age-related memory loss. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 189-203). New York: Oxford University Press.

Parmentier, F. B., King, S., & Dennis, I. (2006). Local temporal distinctiveness does not benefit auditory verbal and spatial serial recall. *Psychonomic Bulletin & Review, 13*, 458-465.

Patterson, M. M., Light, L. L., Van Ocker, J. C., & Olfman, D. (2009). Discriminating semantic from episodic relatedness in young and older adults. *Aging, Neuropsychology, and Cognition, 16*, 535-562.

Perfect, T. J., & Moon, H. C. (2005). The own-age effect in face recognition. In J. Duncan, P. McLeod & L. H. Phillips (Eds.), *Measuring the mind: Speed, control, and age* (pp. 317-340). New York: Oxford University Press.

Perner, J., & Ruffman, T. (1995). Episodic memory and autonoetic consciousness: Developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology*, *59*, 516-548.

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

Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience*, *3*, 85-90.

Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *The Quarterly Journal of Experimental Psychology, 20*, 241-248.

Rabbitt, P. M. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngologica*, *476(Suppl.)*, 167-176.

Raven, J. C., Raven, J., & Court, J. H. (1988). *The Mill Hill vocabulary scale*. London: H. K. Lewis.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 1-90). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15*, 1676-1689.

Rendell, P. G., & Thompson, D. M. (1999). Aging and prospective memory: Differences between naturalistic and laboratory tasks. *Journal of Gerontology, 54B*, P256-P269.

Rhodes, M. G., Castel, A. D., & Jacoby, L. L. (2008). Associative recognition of face pairs by younger and older adults: The role of familiarity-based processing. *Psychology and Aging, 23*, 239-249.

Riley, J. C. (2001). *Rising life expectancy: A global history*. Cambridge: Cambridge University Press.

Roediger III, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

Salthouse, T. A. (1988). Resource reduction interpretations of cognitive aging. *Developmental Review, 8*, 238-272.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging.* Hillsdale: NJ: Lawrence Erlbaum.

Salthouse, T. A. (1993). Speed mediation of adult age differences in cognition. *Developmental Psychology, 29*, 722-738.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403-428.

Salthouse, T. A. (2010). *Major issues in cognitive aging*. New York: Oxford University Press.

Schacter, D. L., Koutstaal, W., & Norman, K. A. (1997). False memories and aging. *Trends in Cognitive Sciences*, *1*, 229-236.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition*, *19*, 523-542.

Schneider, B. A., & Pichora-Fuller, M. K. (2000). Implications of perceptual deterioration for cognitive aging research. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of cognitive aging* (2nd ed., pp. 155-219). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schonfield, D., & Robertson, B. A. (1966). Memory storage and aging. *Canadian Journal of Psychology*, *20*, 228-236.

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Muller, V., Li, S. C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience and Biobehavioral Reviews, 34*, 1080-1091.

Shing, Y. L., Werkle-Bergner, M., Li, S.-C., & Lindenberger, U. (2009). Committing memory errors with high confidence: Older adults do but children don't. *Memory, 17*, 169-179.

Shing, Y. L., Werkle-Bergner, M., Li, S. C., & Lindenberger, U. (2008). Associative and strategic components of episodic memory: A life-span dissociation. *Journal of Experimental Psychology: General, 137*, 495-513.

Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., & Baltes, P. B. (2003). The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychology and Aging, 18,* 318-331.

Skinner, E. I., & Fernandes, M. A. (2008). Interfering with remembering and knowing: Effects of divided attention at retrieval. *Acta Psychologica, 127*, 211-221.

Sluzenski, J., Newcombe, N., & Ottinger, W. (2004). Changes in reality monitoring and episodic memory in early childhood. *Developmental Science, 7*, 225-245.

Sluzenski, J., Newcombe, N. S., & Kovacs, S. L. (2006). Binding, relational memory, and recall of naturalistic events: A developmental perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 89-100.

Smith, A. D., Park, D. C., Earles, J. L., Shaw, R. J., & Whiting, W. L. t. (1998). Age differences in context integration in memory. *Psychology and Aging, 13*, 21-28.

Smith, M. H., Jr. (1948). Spread of the isolation effect in serial learning [Abstract]. *American Psychologist, 3*, 235.

Smith, M. H., Jr. (1949). The influence of isolation on immediate memory. *The American Journal of Psychology, 62*, 405-411.

Smith, M. H., Jr, & Stearns, E. G. (1949). The influence of isolation on the learning of surrounding materials. *The American Journal of Psychology, 62*, 369-381.

Smith, R. E. (2011). Providing support for distinctive processing: The isolation effect in young and older adults. *Psychology and Aging*, Advance online publication. doi: 10.1037/a0022715.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34-50.

Somerville, S. C., & Wellman, H. M. (1979). The development of understanding as an indirect memory strategy. *Journal of Experimental Child Psychology, 27*, 71-86.

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging, 10*, 527-539.

Springer, K., & Murphy, G. L. (1992). Feature availability in conceptual combination. *Psychological Science, 3*, 111-117.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments & Computers, 31*, 137-149.

Stuart, M., Masterson, J., Dixon, M., & Quinlan, P. (1993-1996). Children's printed word database [Data set]. Retrieved from http://www.essex.ac.uk/psychology/cpwd

Treat, N. J., & Reese, H. W. (1976). Age, pacing, and imagery in paired-associate learning. *Developmental Psychology, 12*, 119-124.

Trinkler, I., King, J. A., Spiers, H. J., & Burgess, N. (2006). Part or parcel? Contextual binding of events in episodic memory. In H. D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Handbook of binding and memory:*

Perspectives from cognitive neuroscience (pp. 53-83). New York: Oxford University Press.

Troyer, A. K., & Craik, F. I. M. (2000). The effect of divided attention on memory for items and their context. *Canadian Journal of Experimental Psychology, 54*, 161-170.

Troyer, A. K., Winocur, G., Craik, F. I. M., & Moscovitch, M. (1999). Source memory and divided attention: Reciprocal costs to primary and secondary tasks. *Neuropsychology, 13*, 467-474.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*, 1-12.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-373.

Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76*, 559-573.

Underwood, B. J., & Hughes, R. H. (1950). Gradients of generalized verbal responses. *American Journal of Psychology, 63*, 422-430.

Valentine, T., & Ferrara, A. (1991). Typicality in categorization, recognition and identification: Evidence from face recognition. *British Journal of Psychology, 82*, 87-102.

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging, 18*, 332-339.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1993). Facts and fiction about memory aging: A quantitative integration of research findings. *Journal of Gerontology, 48*, P157-P171.

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology and Aging, 18*, 443-460.

Vitali, P., Minati, L., Chiarenza, G., Brugnolo, A., Girtler, N., Nobili, F., et al. (2006). The von Restorff effect in ageing and Alzheimer's disease. *Neurological Sciences, 27,* 166-172.

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychologische Forschung, 18*, 299-342.

Wallace, W. P. (1965). Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. *Psychological Bulletin, 63,* 410-424.

Wang, W. C., Dew, I. T., & Giovanello, K. S. (2010). Effects of aging and prospective memory on recognition of item and associative information. *Psychology and Aging, 25,* 486-491.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 1*, 442-452.

Wechsler, D. (1981). *Manual for the Wechsler adult intelligence scale – revised.* New York: Psychological Corporation.

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distracters for eyewitness lineups. *Journal of Applied Psychology, 78,* 835-844.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behaviour, 22,* 603-647.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120,* 272-292.

Witte, K. L., Freund, J. S., & Sebby, R. A. (1990). Age differences in free recall and subjective organization. *Psychology and Aging, 5*, 307-309.

Wogalter, M. S., Malpass, R. S., & McQuiston, D. E. (2004). A national survey of US police on preparation and conduct of identification lineups. *Psychology, Crime and Law, 10,* 69-82.

Wright, D. B., & Stroud, J. N. (2002). Age differences in lineup identification accuracy: People are better with their own age. *Law and Human Behaviour, 26*, 641-654.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441-517.

Yoon, C., May, C. P., & Hasher, L. (2000). Aging, circadian arousal patterns, and cognition. In D. C. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 151-171). Philadelphia, PA: Psychology Press.

Zacks, R. T., & Hasher, L. (1997). Cognitive gerontology and attentional inhibition: a reply to Burke and McDowd. *Journal of Gerontology: Psychological Sciences, 52B,* P274-P283.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In T. A. Salthouse & F. I. M. Craik (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 293-357). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Zacks, R. T., Radvansky, G., & Hasher, L. (1996). Studies of directed forgetting in older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 143-156.

Zarkadi, T., Wade, K. A., & Stewart, N. (2009). Creating fair lineups for suspects with distinctive features. *Psychological Science, 20,* 1448-1453.

Zelazo, P. D., Craik, F. I., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica, 115*, 167-183.

Appendix 1

The Digit Symbol Substitution Task

In this task, participants are required to copy a series of symbols as fast as possible into an array of 93 spaces below corresponding digits. The task is presented on a sheet of paper (see Figure i) with a key at the top of the page. The key provides the digits one to nine and below each digit is a corresponding simple shape. Participants are required to use the key to copy the correct shape below each of the digits in the main array. They are given the opportunity to practice seven symbols before completing the main test. During the main test participants are instructed to complete as many of the symbols as possible in 90 s and they are informed that they must complete them in order - from left to right along each row in turn without leaving any gaps. The test is scored as the number of symbols completed correctly in the allowed time.

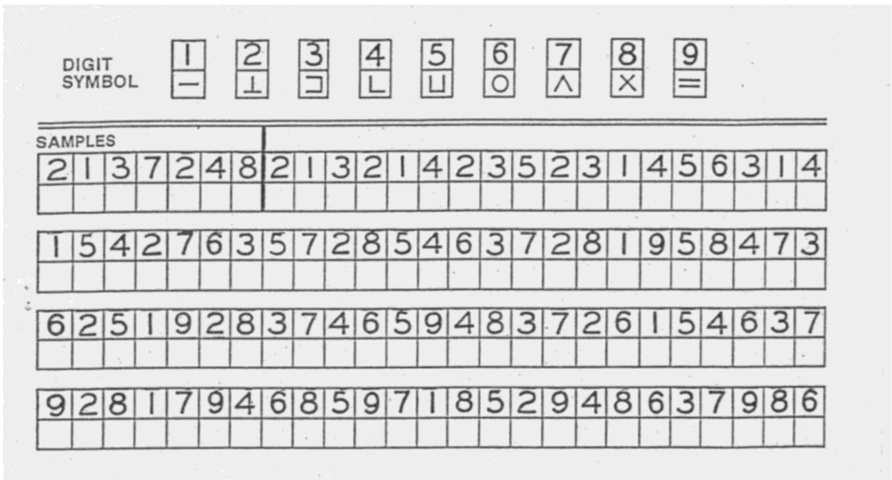

Figure i. Example of a Digit Symbol Substitution Task test sheet.

The Mill Hill Vocabulary Test

In this task, participants are presented with 34 multiple-choice questions on a sheet of paper. Each question has a word in bold and below it six words with a box next to them. Participants must indicate with a tick in the corresponding box which of the six options is closest in meaning to the word at the top of each question. For example, for the top word *brag* the options are: *choose*, *boast*, *hope*, *stone*, *lag* and *jerk*, where *boast* is the correct answer. The first question is completed already as an example; therefore the test is scored out of a maximum of 33. In the current research participants were required to guess if unsure and were given as long as necessary to complete the test.

Appendix 2

Statistical Comparison of Experiment 1 and Bireta et al. (2008)

The data from Experiment 1 and the Bireta et al.'s (2008) fast presentation experiment were entered into a repeated measures ANOVA with age (young, older) and experiment (Experiment 1, Bireta et al. 2008) as between participants factors and with serial position (1-12) and list type (control, isolate) as within participants factors. Mauchly's test indicated that there were violations of sphericity: For serial position, $\chi^2(65) = 607.57$, $p < .001$; therefore, serial position degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .37$). For Serial Position x List type, $\chi^2(65) = 121.40$, $p < .001$; therefore, Serial Position x List type degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .85$).

There was a main effect of age, $F(1, 135) = 66.71$, $MSE = 0.16$, $p < .001$, with young participants remembering more words than older participants on average (*M* (*SD*) = 5.45 (2.52) and 4.08 (2.59), respectively). Surprisingly, there was also a main effect of experiment, $F(1, 135) = 12.18$, $MSE = 0.16$, $p < .001$, with participants from Bireta et al. (2008) remembering more words on average than the participants from Experiment 1 (*M* (*SD*) = 5.06 (2.52) and 4.48 (2.74), respectively). There was a main effect of serial position, $F(4.05, 546.62) = 138.91$, $MSE = 0.14$, $p < .001$, indicating isolation, primacy and recency effects. There was no main effect of list type, $F < 1$. There was an interaction between serial position and age, $F(4.05, 546.62) = 2.47$, $MSE = 0.14$, $p < .05$, driven by a smaller isolation effect in older participants and larger primacy and recency effects in older participants compared to young participants. There was also an interaction between serial position and experiment, $F(4.05, 546.62) = 13.65$, $MSE = 0.14$, $p < .001$; it can be seen in Figures 6 and 7 that there is a larger recency effect in Experiment 1 than in Bireta et al. (2008). There was an interaction between serial position and list type, $F(9.35, 1263.19) = 3.26$, $MSE = 0.03$, $p < .001$, due to the isolation effect. There was also a triple interaction between serial position, list type and experiment, $F(9.35, 1263.19) = 2.11$, $MSE = 0.03$, $p < .05$, driven by a smaller isolation effect in Experiment 1 than in Bireta et al. (2008). None of the other interactions was significant.

Appendix 3

Nonwords Used in Experiment 4

abblaims	chrimmage	glockdown	prallpox	sproolboy
addaches	cleatment	glofold	prapboard	stowtorch
aggairs	clindling	glonewall	preepest	thordplay
althan	cloubled	glorecard	puddlong	thowshoe
annepts	cracesuit	granksep	pulholes	threaker
appess	crallest	haurded	sangaurd	threndid
attressed	drassics	laurdroom	scheesome	todblar
bandwidsk	drinflint	nogwhell	scrowback	tossens
bartlech	drirsted	noncieves	shagrant	tranders
bleersman	dritfire	ollshoot	slactised	treelyard
bligma	droneware	ottress	slanquil	trillborn
brooklum	drowflake	peptokes	slansoms	trockpile
broonfeed	flinness	pheties	sleshmen	troodless
brushfute	forthwirt	plarkled	sminkles	weathezod
chakeout	fralltime	pleepings	snushwork	whiketead
chanslate	frunning	pligots	spavel	wookyurd
chipod	gentips	plimming	spleetcar	wortsmeet
chitching	gleatband	plinhead	splotum	yorpsmire

Appendix 4

Stimuli Used in Experiments 5 and 6

Prime/Cue			
Integrative	Semantic	Unrelated (Exp. 6 only)	Target
travel	article	lapel	book
lemon	muffin	affection	cake
soup	jug	stable	can
birthday	flashlight	pillow	candle
race	motorcycle	author	car
town	convent	athlete	church
necklace	pearl	stick	diamond
horse	sick	pub	doctor
apartment	fox	company	dog
velvet	lady	cow	dress
ocean	lobster	guide	fish
monkey	paw	campus	foot
herb	lawn	towel	garden
halloween	vampire	celebration	ghost
jelly	cherry	fence	grape
donor	liver	icing	heart
brass	clarinet	light	horn
parade	ox	theory	horse
beach	palace	mushroom	house
thesis	insight	fall	idea
border	field	party	land
maple	branch	valentine	leaf
government	fact	flower	lie
puppy	trust	pool	love
deer	vegetable	umbrella	meat
strawberry	juice	plumber	milk
copper	credit	carrot	money
farm	chipmunk	stairway	mouse
linen	blouse	estuary	pants
rice	envelope	gear	paper
concert	harp	square	piano
steel	tube	fight	pipe
corporate	rocket	plug	plane
trick	mole	industry	rabbit
summer	tornado	food	rain
law	office	acre	school
airplane	fatigue	glass	sleep
jungle	crocodile	hat	snake
mountain	wind	wick	snow
bathroom	shampoo	island	soap
winter	tennis	termite	sport
gold	tongue	lecture	teeth
plastic	game	smoke	toy
box	gin	remote	wine
fireplace	coal	chain	wood

Appendix 5

The Seven Categories Based on Semantic Themes and Corresponding Words Used in the Study Period and the New Words Used in the Item Memory Test (Experiment 7)

Semantic Category	Words Used
Weather/Seasons	rain, storm, summer, spring, snow, winter, autumn, sun, clouds, sky, weather, wind
Water	water, drop, pool, river, stream, sea, shore, well, sand, lake, damp, bridge
Animals	cat, dog, chicken, hen, lion, tiger, duck, monkey, fish, kitten, mouse, rat
Body Parts	head, neck, feet, hands, arm, leg, mouth, ears, hair, nose, tummy, eyes
Nature	tree, branch, leaf, flower, daisy, forest, woods, hedge, bush, field, grass, rose
People	man, lady, girl, boy, mum, dad, woman, queen, king, aunt, son
Time	hour, minute, time, second, clock, watch, wrist, morning, night, day, months, year
New words used in item test	air, apple, balloon, ball, bath, chair, cheese, clothes, doctor, dragon, feast, fruit, gold, hat, heart, ink, jungle, letter, map, oil, paint, pirate, radio, shop, sign, toilet, torch, window

Appendix 6

Item and Associative Recognition Memory Performance and Response Bias for each Divided Attention Condition and Test Order (Experiment 8)

DA Condition	DA-FA order	Attention	Item H-FA		Assoc H-FA		Item d'		Assoc d'		Item lnβ		Assoc lnβ	
			M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
DA-E	DA-FA	DA	0.49	0.24	0.28	0.36	1.51	0.86	0.84	1.06	0.03	0.74	0.18	0.34
		FA	0.55	0.28	0.53	0.24	1.82	1.15	1.64	0.86	-0.12	0.71	0.02	0.63
	FA-DA	DA	0.50	0.26	0.52	0.24	1.59	0.98	1.64	0.92	0.29	0.61	0.26	0.56
		FA	0.70	0.19	0.72	0.28	2.46	0.83	2.48	1.08	0.62	0.94	0.15	0.41
DA-R	DA-FA	DA	0.58	0.19	0.35	0.30	1.86	0.72	1.02	0.93	0.52	0.70	0.16	0.44
		FA	0.58	0.23	0.55	0.33	1.87	0.88	1.69	1.08	0.60	0.48	0.26	0.63
	FA-DA	DA	0.58	0.20	0.49	0.29	1.88	0.80	1.50	0.98	0.56	0.77	0.17	0.45
		FA	0.73	0.10	0.62	0.29	2.37	0.51	1.91	0.98	0.63	0.64	-0.13	0.67
DA-E&R	DA-FA	DA	0.42	0.29	0.35	0.36	1.37	1.00	1.10	1.18	0.40	0.61	0.16	0.46
		FA	0.61	0.24	0.45	0.38	2.00	0.99	1.53	1.34	-0.15	0.66	0.06	0.53
	FA-DA	DA	0.49	0.19	0.34	0.23	1.48	0.59	1.02	0.84	0.03	0.59	0.22	0.30
		FA	0.68	0.16	0.57	0.33	2.24	0.72	1.88	1.22	0.44	0.81	0.08	0.57

Note. Attention: Divided attention (DA), full attention (FA). Hits minus false alarms (H-FA). DA type: DA at encoding only (DA-E), DA at retrieval only (DA-R), DA at both encoding and retrieval (DA-E&R). DA test order: DA test before FA test (DA-FA), DA test after FA test (FA-DA).

End